TRANSACTIONS

OF THE

AMERICAN MATHEMATICAL SOCIETY

EDITED BY

ROBERT D. CARMICHAEL

WILLIAM G. GRAUSTEIN

JACOB D. TAMARKIN

WITH THE COÖPERATION OF

A. A. ALBERT	ERIC T. BELL	JESSE DOUGLAS
T. H. HILDEBRANDT	EINAR HILLE	E. P. LANE
R. E. LANGER	C. C. MACDUFFEE	MARSTON MORSE
G. Y. RAINICH	H. L. RIETZ	M. H. STONE
J. L. SYNGE	G. T. WHYBURN	OSCAR ZARISKI

VOLUME 40

JULY TO DECEMBER, 1936

PUBLISHED BY THE SOCIETY

MENASHA, WIS., AND NEW YORK

1936

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

General
Bound March 1937

Composed, Printed and Bound by
The Collegiate Press
George Banta Publishing Company
Menasha, Wisconsin

55939

22

QA
1
F001
v.40

TABLE OF CONTENTS
VOLUME 40, JULY TO DECEMBER, 1936

ADAMS, C. R., of Providence, R. I. The space of functions of bounded variation and certain general spaces 421
ALBERT, A. A., of Chicago, Ill. Simple algebras of degree p^e over a centrum of characteristic p 112
BIRD, M. T., CARMICHAEL, R. D., and MARTIN, W. T., of Urbana, Ill. On a classification of integral functions by means of certain invariant point properties 462
BOAS, R. P., of Cambridge, Mass. Some theorems on Fourier transforms and conjugate trigonometric integrals 287
BOCHNER, S., of Princeton, N. J. Summation of multiple Fourier series by spherical means 175
CAIRNS, S. S., of Bethlehem, Pa. The generalized theorem of Stokes 167
CARMICHAEL, R. D., MARTIN, W. T., and BIRD, M. T., of Urbana, Ill. On a classification of integral functions by means of certain invariant point properties 462
CLARKSON, J. A., of Princeton, N. J. Uniformly convex spaces 396
DUNFORD, NELSON, of Providence, R. I. Integration and linear operations 474
DUNFORD, N., and MORSE, A. P., of Providence, R. I. Remarks on the preceding paper of James A. Clarkson 415
FIALKOW, A., of New York, N. Y. Initial motion in fields of force 495
GROVE, V. G., of East Lansing, Mich. Differential geometry of a surface at a planar point 155
HULL, R., of Chicago, Ill. The maximal orders of generalized quaternion division algebras 1
JACKSON, D., of Minneapolis, Minn. Bernstein's theorem and trigonometric approximation 225
KETCHUM, G. S., of Urbana, Ill. On certain generalizations of the Cauchy-Taylor expansion theory 208
LATIMER, C. G., of Lexington, Ky. On the class number of a quaternion algebra with a negative fundamental number 318
——— On ideals in a quaternion algebra and the representation of integers by Hermitian forms 439
LEIGHTON, W., and MORSE, M., of Cambridge, Mass. Singular quadratic functionals 252
MCKINSEY, J. C. C., of Berkeley, Calif. On Boolean functions of many variables 343
MACLANE, S., of Cambridge, Mass. A construction for absolute values in polynomial rings 363

MANNING, D., of Palo Alto, Calif. On simply transitive groups with transitive Abelian subgroups of the same degree 324
MARTIN, W. T., BIRD, M. T., and CARMICHAEL, R. D., of Urbana, Ill. On a classification of integral functions by means of certain invariant point properties 462
MONTGOMERY, D., and ZIPPIN, L., of Princeton, N. J. Periodic one-parameter groups in three-space 24
MORSE, A. P., and DUNFORD, N., of Providence, R. I. Remarks on the preceding paper of James A. Clarkson 415
MORSE, M., and LEIGHTON, W., of Cambridge, Mass. Singular quadratic functionals 252
SCHEFFÉ, H., of Madison, Wis. Asymptotic solutions of certain linear differential equations in which the coefficient of the parameter may have a zero 127
SCHOENBERG, I. J., of Swarthmore, Pa. On the zeros of successive derivatives of integral functions 12
STONE, M. H., of Cambridge, Mass. The theory of representations for Boolean algebras 37
SZEGÖ, G., of St. Louis, Mo. On some Hermitian forms associated with two given curves of the complex plane 450
WHITNEY, H., of Cambridge, Mass. Differentiable functions defined in arbitrary subsets of Euclidean space 309
ZIPPIN, L., and MONTGOMERY, D., of Princeton, N. J. Periodic one-parameter groups in three-space 24

THE MAXIMAL ORDERS OF GENERALIZED QUATERNION DIVISION ALGEBRAS*

BY

RALPH HULL†

1. **Introduction.** A generalized quaternion division algebra Q is an algebra of degree 2, order 4, over the field R of all rational numbers. It can be written in the form $Q=(a, Z)=Z+uZ$, where Z is a quadratic field over R and $u^2=a$ is in R and is not the norm of an element of Z, and $Zu=uZ'$, elementwise, where the prime denotes the conjugate in Z. We say that Q has the generation $Q=(a, Z)$ and identify the unity elements‡ of Q and Z with each other and with 1, the unity element of R. We call 1, u, a Z-basis of Q. If ζ_1, ζ_2, and ζ_3, ζ_4, are any R-bases of Z then $\zeta_1, \zeta_2, u\zeta_3, u\zeta_4$, is an R-basis of Q.

An algebra Q has a representation as an algebra of matrices of degree 2 with elements in Z which can be obtained in the following manner. Regarding $(1, u)$ as a vector, for any $q=\zeta_0+u\zeta_1$ of Q, ζ_0 and ζ_1 in Z, we have $q \longleftrightarrow \bar{q}$, where

$$q(1, u) = (q, qu) = (1, u)\bar{q}, \quad \bar{q} = \begin{Vmatrix} \zeta_0, & a\zeta_1' \\ \zeta_1, & \zeta_0' \end{Vmatrix}.$$

We call $T(\bar{q})=T(q)=\zeta_0+\zeta_1$ the *reduced trace* of q. In a similar way, using an R-basis of Q we obtain a representation of Q as an algebra of matrices of degree 4 with elements in R.

The Q are cyclic algebras§ and the theory of their invariants is included in the general theory of Hasse‖ which yields all generations of a given Q. There exist a finite number >0 of rational primes π such that the π-adic extension Q_π of Q is a division algebra whereas Q_ρ for all other rational primes ρ is a total matric algebra. We say that Q splits at the prime spots ρ of R. The number of π is even if Q has a real quadratic sub-field, and odd otherwise. In the former case, Q is said to split at the infinite prime spot of R.

A *maximal order*¶ $\mathfrak{M}$ of Q, that is, an integral domain (Dickson, loc. cit.,

* Presented to the Society, December 31, 1935; received by the editors December 2, 1935.

† The paper had its inception while its author was a National Research Fellow.

‡ The name *unity element* in preference to *modulus* or *unit element* was adopted at the suggestion of Professor A. A. Albert.

§ Dickson, *Algebren und ihre Zahlentheorie*.

‖ Hasse, *Theory of cyclic algebras over an algebraic number field*, these Transactions, vol. 34 (1932), pp. 171–214.

¶ Deuring, *Algebren*, Ergebnisse der Mathematik, vol. 4, part 1. Deuring gives a complete bibliography including references to the work of Darkow and Latimer.

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

p. 198), is a set of elements of Q with the properties:

U: $\mathfrak{M}$ contains the unity element.

B: $\mathfrak{M}$ contains an R-basis of Q.

I: The elements of $\mathfrak{M}$ are integral; that is, they satisfy equations with rational integral coefficients, highest coefficient 1.

C_a: $\mathfrak{M}$ is closed under addition.

C_m: $\mathfrak{M}$ is closed under multiplication.

M: $\mathfrak{M}$ is maximal; that is, it is not contained in a larger set having the first properties.

A set having properties $U, \cdots, C_m$, is called an *order* $\mathfrak{R}$ of Q.

Each Q has infinitely many $\mathfrak{M}$ of which special ones have been determined for certain Q by Dickson (loc. cit.), Darkow and Latimer, and for all Q by Albert.* It is the purpose of this paper to determine all $\mathfrak{M}$ for each Q.

First, every $\mathfrak{R}$ is shown to have a certain simple form relative to an arbitrary generation of Q. Second, the necessary and sufficient conditions for a set of elements of this form to be a $\mathfrak{R}$ are determined. Third, the maximality condition is introduced. Fourth, canonical generations of each Q found by Albert (loc. cit.) are described. Finally, conditions for the existence of maximal orders $\mathfrak{M}$ are determined by a study of the earlier results when expressed in terms of canonical generations.

Throughout the paper we denote by $\mathfrak{g}$ the maximal order of R. It is clear by Properties U and C_a that every $\mathfrak{R}$ contains $\mathfrak{g}$. Also, if v_1 and v_2 are fixed quantities and $\mathfrak{m}$ and $\mathfrak{n}$ are sets of quantities we write $v_1 \cdot \mathfrak{m} + v_2 \cdot \mathfrak{n}$ for the set of all quantities of the form $v_1\mu + v_2\nu$, μ in $\mathfrak{m}$ and ν in $\mathfrak{n}$.

2. **The form of an order.** Any order $\mathfrak{R}$ of an algebra Q has the simple form relative to an arbitrary generation $Q = (a, Z)$, described in

THEOREM 1. *If $\mathfrak{R}$ is any order of $Q = (a, Z)$ the intersection $\mathfrak{m}_c$ of $\mathfrak{R}$ and Z is an order of Z whose conductor is a positive integer c. There exists a unique $\mathfrak{m}_c$-modul $\mathfrak{n}$, which is a finite $\mathfrak{g}$-modul, of elements of Z, and a quantity λ of Z such that*

$$\mathfrak{R} = 1 \cdot \mathfrak{m}_c + (\lambda + u) \cdot \mathfrak{n}. \tag{1}$$

By the definition of $\mathfrak{m}_c$ and the order properties of $\mathfrak{R}$ it is evident that $\mathfrak{m}_c$ is an order of Z. It is known that the conductor of every order of a quadratic field is a positive integer c, uniquely determining the order, such that $c^2 d$ is the discriminant of the order, where d is the discriminant of the field.

By Property B, $\mathfrak{R}$ contains elements of the form $\nu_0 + u\nu$, $\nu \neq 0$ and ν_0 in Z. Let $\mathfrak{n}$ be the set of all ν, including $\nu = 0$, which occur when all elements of $\mathfrak{R}$ are

* Albert, *Integral domains of rational generalized quaternion algebras*, Bulletin of the American Mathematical Society, vol. 40 (1934), pp. 164–176.

written in this form. By Property C_m, $\mathfrak{R}\mathfrak{m}_c \subset \mathfrak{R}$ whence $\mathfrak{n}\mathfrak{m}_c \subset \mathfrak{n}$. Hence $\mathfrak{n}$ is an $\mathfrak{m}_c$-modul since Property C_a of $\mathfrak{R}$ implies that $\mathfrak{n}$ is a modul.

To prove that $\mathfrak{n}$ is a finite $\mathfrak{g}$-modul we prove more, namely, that $\mathfrak{R}$ is a finite $\mathfrak{g}$-modul. For, $\mathfrak{R}$ contains an R-basis, say v_0, v_1, v_2, v_3, of Q by Property B. Let q be any element of $\mathfrak{R}$. Then we have

$$q = \sum_{i=0}^{3} \alpha_i v_i, \qquad \alpha_i \text{ in } R, \qquad qv_j = \sum_{i=0}^{3} \alpha_i v_i v_j \qquad (j = 0, \cdots, 3).$$

From these equations we get

$$T(qv_j) = \sum_{i=0}^{3} \alpha_i T(v_i v_j) \qquad (j = 0, \cdots, 3),$$

where all traces are in $\mathfrak{g}$ by Properties C_m and I. Solution shows that $|T(v_i v_j)|\alpha_k$ is in $\mathfrak{g}$ for $k=0, \cdots, 3$, where $|T(v_i v_j)| = \Delta(v_0, \cdots, v_3)$ is in $\mathfrak{g}$, is independent of q in $\mathfrak{R}$ and is not zero since Q is semi-simple. This proves that $\mathfrak{R}$ is a finite $\mathfrak{g}$-modul and it is evident, therefore, that $\mathfrak{n}$ is a finite $\mathfrak{g}$-modul. We call $\Delta(v_0, \cdots, v_3)$ the *reduced discriminant* of the R-basis $v_0, \cdots, v_3$.

From the properties of $\mathfrak{n}$ just proved it follows in the usual way that there exist $\nu_1 \neq 0$ and $\nu_2 \neq 0$ in $\mathfrak{n}$ such that $\mathfrak{n} = \nu_1 \cdot \mathfrak{g} + \nu_2 \cdot \mathfrak{g}$. Then $\mathfrak{R}$ contains quantities $\nu_{01} + u\nu_1$, $\nu_{02} + u\nu_2$, ν_{01} and ν_{02} in Z. By Property C_m, $\mathfrak{R}$ also contains

$$(\nu_{01} + u\nu_1)(\nu_{02} + u\nu_2) = \nu_{01}\nu_{02} + a\nu_1'\nu_2 + u(\nu_1\nu_{02} + \nu_2\nu_{01}'),$$

whence $\nu_1\nu_{02} + \nu_2\nu_{01}'$ is in $\mathfrak{n}$. By Property I,

$$T(\nu_{01} + u\nu_1) = \nu_{01} + \nu_{01}' = g$$

is in $\mathfrak{g}$. Hence

$$\nu_1\nu_{02} - \nu_2\nu_{01} = \nu_1\nu_{02} + \nu_2\nu_{01}' - g\nu_2$$

is in $\mathfrak{n}$ and there exist g_1 and g_2 in $\mathfrak{g}$ such that

$$\nu_1\nu_{02} - \nu_2\nu_{01} = g_2\nu_1 - g_1\nu_2, \qquad (\nu_{01} - g_1)\nu_2 = (\nu_{02} - g_2)\nu_1.$$

We define $\lambda = (\nu_{01} - g_1)/\nu_1 = (\nu_{02} - g_2)/\nu_2$ and prove (1).

We have $(\lambda + u)\nu_1 = \nu_{01} + u\nu_1 - g_1$ in $\mathfrak{R}$ by Property C_a since $\mathfrak{g} \subset \mathfrak{R}$. Similarly, $(\lambda + u)\nu_2$ is in $\mathfrak{R}$ and hence by C_a, $(\lambda + u)\mathfrak{n} \subset \mathfrak{R}$. For an arbitrary $\nu_0 + u\nu$ in $\mathfrak{R}$, whence ν is in $\mathfrak{n}$, we have $\nu_0 + u\nu = \mu + (\lambda + u)\nu$, $\mu = \nu_0 - \lambda\nu$, where μ is in Z and also in $\mathfrak{R}$ by Property C_a and hence in $\mathfrak{m}_c$. Since $\mathfrak{m}_c \subset \mathfrak{R}$ and $(\lambda + u)\mathfrak{n} \subset \mathfrak{R}$, we have (1).

An order $\mathfrak{m}_c$ of Z and a finite $\mathfrak{g}$-modul $\mathfrak{n}$ of elements of Z such that $\mathfrak{n}\mathfrak{m}_c \subset \mathfrak{n}$, have $\mathfrak{g}$-bases of a special form needed later. We state*

* For the first part of the theorem, see Fricke, *Lehrbuch der Algebra*, vol. 3, p. 249. The second part is easily proved as a consequence of the properties of $\mathfrak{n}$.

THEOREM 2. *The order* $\mathfrak{m}_c$ *of* Z, *with the conductor* c, *can be written*

$$\mathfrak{m}_c = 1\cdot\mathfrak{g} + c\omega\cdot\mathfrak{g}, \qquad \omega = (d + d^{1/2})/2, \tag{2}$$

where d *is the discriminant of* Z. *A finite* $\mathfrak{g}$*-modul* $\mathfrak{n}$ *of elements of* Z *such that* $\mathfrak{n}\mathfrak{m}_c \subset \mathfrak{n}$, *can be written*

$$\mathfrak{n} = r\cdot\mathfrak{g} + rv\cdot\mathfrak{g}, \qquad v = (g_1 + c\omega)/g_2, \tag{3}$$

where r *is in* R *and* g_1 *and* g_2 *are in* $\mathfrak{g}$ *and such that*

$$g_1^2 - cdg_1 + c^2(d^2 - d)/4 \equiv 0 \pmod{g_2}. \tag{4}$$

3. **The closure and integral conditions.** Let $\mathfrak{m}_c$ and $\mathfrak{n}$ be given by (2), (3), and (4). Let

$$\lambda = e_0 + e_1\omega, \qquad e_0 \text{ and } e_1 \text{ in } R, \tag{5}$$

whence λ is in Z. We consider the set

$$\mathfrak{S} = 1\cdot\mathfrak{m}_c + (\lambda + u)\cdot\mathfrak{n}, \tag{6}$$

of elements of Q. Evidently $\mathfrak{S}$ has Properties U, C_a, and B, since 1 is in $\mathfrak{m}_c$, $\mathfrak{m}_c$ and $\mathfrak{n}$ have Property C_a, and v_0, v_1, v_2, v_3, where

$$v_0 = 1, \qquad v_1 = c\omega, \qquad v_2 = (\lambda + u)r, \qquad v_3 = (\lambda + u)rv, \tag{7}$$

is an R-basis of Q in $\mathfrak{S}$. We shall determine necessary and sufficient conditions that $\mathfrak{S}$ have Properties I and C_m, and hence be an order of Q. We first prove

LEMMA 1. *The trace* $T(s)$ *is in* $\mathfrak{g}$ *for every* s *of* S *if and only if* $T(\lambda r) = k_1$ *and* $T(\lambda rv) = k_2$ *are in* $\mathfrak{g}$. *These conditions are equivalent to*

$$\begin{aligned} rce_0 &= \{g_1 + c(d + 1)/2\}k_1 - g_2k_2, \\ rcde_1 &= -(2g_1 + cd)k_1 + 2g_2k_2, \end{aligned} \tag{8}$$

with k_1 *and* k_2 *in* $\mathfrak{g}$.

If $s = \mu + (\lambda + u)\eta$ is in $\mathfrak{S}$, we have μ in $\mathfrak{m}_c$ and η in $\mathfrak{n}$. Then $T(\mu)$ is in $\mathfrak{g}$ and $T(s) = T(\mu) + T(\lambda\eta)$ is in $\mathfrak{g}$ if and only if $T(\lambda\eta)$ is in $\mathfrak{g}$. From (3) and the linearity of the trace function we obtain at once the lemma, where (8) is the solution of $T(\lambda r) = k_1$ and $T(\lambda rv) = k_2$ for e_0 and e_1.

We leave Property I and consider C_m assuming (2), $\cdots$, (8). Since $\mathfrak{m}_c\cdot\mathfrak{m}_c = \mathfrak{m}_c$ and $\mathfrak{n}\cdot\mathfrak{m}_c = \mathfrak{n}$, it follows from (6) that $\mathfrak{S}$ has Property C_m if and only if we have simultaneously:

$$\mu(\lambda + u)\eta \text{ in } \mathfrak{S} \text{ for every } \mu \text{ in } \mathfrak{m}_c, \eta \text{ in } \mathfrak{n}, \tag{9}$$

and

$$(\lambda + u)\eta_1(\lambda + u)\eta_2 \text{ in } \mathfrak{S} \text{ for every } \eta_1, \eta_2 \text{ in } \mathfrak{n}. \tag{10}$$

By (2) and (3) it is necessary and sufficient for (9) that $c\omega(\lambda+u)r$ and $c\omega(\lambda+u)r\nu$ be in $\mathfrak{S}$. Using (2), $\cdots$, (8) we find

$$
\begin{aligned}
c\omega(\lambda+u)r &= -(g_1+cd)k_1+g_2k_2+k_1c\omega+(\lambda+u)c\omega' r,\\
c\omega(\lambda+u)r\nu &= -g_3k_1+g_1k_2+k_1c\omega+(\lambda+u)c\omega' r,
\end{aligned}
\tag{11}
$$

where $g_3g_2=g_1^2+cdg_1+c^2(d^2-d)/4$, whence g_3 is in $\mathfrak{g}$ by (4). Since $c\omega'$ is in $\mathfrak{m}_c$, $c\omega' r$ and $c\omega' r\nu$ are in $\mathfrak{n}$. Hence (11) shows that (2), $\cdots$, (8) imply (9).

Next, we have

$$
(\lambda+u)\eta_1(\lambda+u)\eta_2=\{a-N(\lambda)\}\eta_1'\eta_2+(\lambda+u)\eta_2T(\lambda\eta_1),
\tag{12}
$$

where $N(\lambda)=\lambda\lambda'$. For every η_1, η_2 in $\mathfrak{n}$, $T(\lambda\eta_1)$ is in $\mathfrak{g}$ by Lemma 1 and hence $\eta_2T(\lambda\eta_1)$ is in $\mathfrak{n}$. Hence, by (3), (4), and (12) it is necessary and sufficient for (10) that $\{a-N(\lambda)\}\eta_1'\eta_2$ be in $\mathfrak{m}_c$ in the four cases: $\eta_1=\eta_2=r$; $\eta_2=r\nu$; $\eta_1=r\nu$, $\eta_2=r$; $\eta_1=\eta_2=r\nu$. In the second of these cases (10) requires that $r^2\{a-N(\lambda)\}\nu$ be in $\mathfrak{m}_c$ which holds, by (3), if and only if

$$
r^2\{a-N(\lambda)\}=k_3g_2,\qquad k_3 \text{ in } \mathfrak{g}.
\tag{13}
$$

It is readily seen that the first and third cases require no additional conditions. In view of (4), (13) is also necessary and sufficient in the fourth case. This completes the proof of

LEMMA 2. *If the elements of $\mathfrak{S}$ have integral traces, $\mathfrak{S}$ has Property C_m if and only if* (13) *holds.*

We now return to Property *I* and prove

LEMMA 3. *If $\mathfrak{S}$ has Property C_m then $\mathfrak{S}$ has Property I.*

It is plain that (7) is a $\mathfrak{g}$-basis of $\mathfrak{S}$. For any s in $\mathfrak{S}$ we have

$$
s=\sum_{i=0}^{3}\sigma_iv_i,\qquad \sigma_i \text{ in } \mathfrak{g}\qquad (i=0,\cdots,3).
$$

By means of this basis, in the manner indicated in the Introduction, we form the representation of Q as an algebra of matrices of degree 4 with elements in R. In this representation s in $\mathfrak{S}$ corresponds to a matrix with elements in $\mathfrak{g}$ by Property C_m. Hence s is integral since it satisfies the characteristic equation of this matrix.

Combining Lemmas 1, 2, and 3 we have

THEOREM 3. *The set $\mathfrak{S}$ in* (6) *is an order of Q if and only if* (8) *and* (13) *hold with k_1, k_2, and k_3 in $\mathfrak{g}$.*

4. **The maximality condition.** We now study Property *M* in connection with the reduced discriminant, for brevity, discriminant, of an order.

This is defined* for algebras Q over R, as the discriminant of any $\mathfrak{g}$-basis (see §2) of the order. By the following lemma, the discriminant of an order is independent of the $\mathfrak{g}$-basis.

LEMMA. *If* $v_0, \cdots, v_3$, *and* $u_0, \cdots, u_3$, *are two* R*-bases of* Q, *such that* $(u_0, \cdots, u_3) = A(v_0, \cdots, v_3)$, *where* A *is a rational matrix of degree* 4, *then*

$$\Delta(u_0, \cdots, u_3) = |A|^2 \Delta(v_0, \cdots, v_3).$$

The lemma is a well known consequence of the linearity and symmetry of the trace function. If the u's and v's are $\mathfrak{g}$-bases of the same order, $|A|^2 = 1$.

The discriminant of $\mathfrak{R}$ in Theorems 1 and 2 is

$$\Delta(\mathfrak{R}) = -r^4 c^4 a^2 d^2 / g_2^2 . \tag{14}$$

This can be verified by computing the discriminant of the $\mathfrak{g}$-basis (7) of $\mathfrak{R}$ by means of the lemma from $\Delta(1, c\omega, u, uc\omega)$ which is easily found directly.

The discriminant of a maximal order $\mathfrak{M}$ of Q, which is invariant for all $\mathfrak{M}$, is called the discriminant $\Delta(Q)$ of Q. It is known (cf. Reichardt, loc. cit.) that π^2 divides $\Delta(Q)$, π^3 does not divide $\Delta(Q)$, for each of the primes π described in the Introduction, and that $\Delta(Q)$ is not divisible by any other rational prime. For the purposes of the next section it is convenient to define a_0 as the product of the π or as the negative of their product according as the number of π is even or odd. With (14) and the lemma these remarks imply $\Delta(Q) = -a_0^2$ and yield

THEOREM 4. *An order of* Q *given by Theorems* 1, 2, *and* 3 *is maximal if and only if*

$$r^4 c^4 a^2 d^2 = a_0^2 g_2^2 . \tag{15}$$

5. **Canonical generations.** We now describe canonical generations of the Q in

THEOREM 5. *Each* Q *has a canonical generation* $Q = (a_0, P)$, *where* a_0 *is the quantity defined in* §4 *and* P *is any quadratic field with the following properties. The discriminant of* P *is* $-p$, *where* p *is a prime such that* $p \equiv 3 \pmod 4$ *and each prime factor of* a_0 *is a quadratic non-residue, while* a_0 *is a quadratic residue modulo* p. *The discriminant of* Q *is* $-a_0^2$.

This theorem follows at once from theorems of Albert (loc. cit., Theorems 1, 2, and 3) by verifying, either by the use of Hasse's theory, or by computing the discriminant of the special maximal orders found by Albert, that the

* See Reichardt, *Die Diskriminante einer normalen einfachen Algebra*, Journal für Mathematik, vol. 172 (1935), pp. 31–35, for the extension of the definition to the case of an algebra over an algebraic number field and for the form of the discriminant in terms of ideals of the coefficient field.

quantity σ of Albert's Theorem 3 is the quantity a_0 we defined in §4. Similar canonical generations exist* for cyclic division algebras of odd prime degree over R.

To express the conditions of Theorems 1, $\cdots$, 4 in terms of a canonical generation we replace a by a_0 and d by $-p$. We assume, without loss of generality, that $g>0$ and $r>0$. Then from (15) we get $g_2^2 = r^4c^4p^2$, $g_2 = r^2c^2p$, whence $g_2 = pg^2$, g in $\mathfrak{g}$, $r = g/c$. Next, (4) requires $g_1 = ph$, h in $\mathfrak{g}$, and

$$ph^2 - cph + c^2(p+1)/4 = g_3g^2 \equiv 0 \pmod{g^2}. \tag{16}$$

From (5) and (8), we now obtain

$$\begin{aligned} \lambda &= e_0 + e_1\omega, \\ ge_0 &= \{ph - c(p-1)/2\}k_1 - pg^2k_2, \\ ge_1 &= (2h - c)k_1 - 2g^2k_2. \end{aligned} \tag{17}$$

Finally, we compute $N(\lambda) = \lambda\lambda'$ from (17) and substitute in (13) which becomes

$$a_0 - g_3k_1^2 + p(2h - c)k_1k_2 - pg^2k_2^2 = k_3c^2p \equiv 0 \pmod{c^2p}. \tag{18}$$

This completes the proof of

THEOREM 6. *Let $Q = (a_0, P)$ be a canonical generation. Every maximal order $\mathfrak{M}$ of Q is of the form*

$$\mathfrak{M} = 1\cdot\mathfrak{m}_c + (\lambda + u)\mathfrak{n}, \tag{19}$$

where

$$\mathfrak{m}_c = 1\cdot\mathfrak{g} + c\omega\cdot\mathfrak{g}, \qquad \omega = \{-p + (-p)^{1/2}\}/2, \qquad c \geqq 1 \text{ in } \mathfrak{g}, \tag{20}$$

is the intersection of $\mathfrak{M}$ and P;

$$\mathfrak{n} = r\cdot\mathfrak{g} + rv\cdot\mathfrak{g}, \qquad r = g/c, \qquad v = (ph + c\omega)/pg^2, \tag{21}$$

where g and h are in $\mathfrak{g}$ and satisfy (16), *and λ in Z is given by* (17) *for k_1, and k_2 in $\mathfrak{g}$ such that* (18) *holds. Conversely, if c, g, and h are such that* (16) *holds and there exists a solution k_1, k_2 in $\mathfrak{g}$ of* (18), *define $\mathfrak{m}_c$ and $\mathfrak{n}$ by* (20) *and* (21). *Then if λ is given by* (17) *with any solution of* (18), *$\mathfrak{M}$ in* (19) *is a maximal order of Q whose intersection with P is $\mathfrak{m}_c$.*

The question as to the existence of maximal orders $\mathfrak{M}$, that is, the existence of integers c, g, h, etc., satisfying the conditions of Theorem 6, is deferred to the next section. For the sake of completeness we here adjoin the multiplication table of a $\mathfrak{g}$-basis of $\mathfrak{M}$ in (19). No use is made of this in what follows. We have

* Cf. Hull, these Transactions, vol. 38 (1935), p. 517.

$$\mathfrak{M} = v_0 \cdot \mathfrak{g} + v_1 \cdot \mathfrak{g} + v_2 \cdot \mathfrak{g} + v_3 \cdot \mathfrak{g},$$
$$v_0 = 1, \qquad v_1 = c\omega, \qquad v_2 = (\lambda + u)r, \qquad v_3 = (\lambda + u)rv,$$

where all quantities are as defined in Theorem 6. The following relations hold:

$$v_0 v_i = v_i v_0 = v_i \qquad (i = 0, \cdots, 3)$$
$$v_1^2 = -c^2(p^2 + p)/4 - cpv_1,$$
$$v_1 v_2 = pg^2 k_2 - p(h - c)k_1 + k_1 v_1 + p(h - c)v_2 - pg^2 v_3,$$
$$v_1 v_3 = hpk_2 - g_3 k_1 + k_2 v_1 + g_3 v_2 - hpv_3,$$
$$v_2 v_1 = -hpv_2 + pg^2 v_3, \qquad v_2^2 = pg^2 k_3 + k_1 v_2,$$
$$v_2 v_3 = hpk_3 + k_3 v_1 + k_1 v_3, \qquad v_3 v_1 = -g_3 v_2 + p(h - c)v_3,$$
$$v_3 v_2 = p(h - c)k_3 - k_3 v_1 + k_2 v_2, \qquad v_3^2 = g_3 k_3 + k_2 v_3.$$

6. **On the existence of maximal orders.** As a special case of a general theorem of Schilling* on division algebras of prime degree over an algebraic field, we have the following existence theorem.

THEOREM 7. *There exists a maximal order of $Q = (a_0, P)$ whose intersection with P is a given order $\mathfrak{m}_c$ of P if and only if c is prime to the discriminant of Q, that is, if and only if*

$$(c, a_0) = 1. \tag{22}$$

This theorem can also be proved directly by means of Theorems 5 and 6. We shall indicate a proof of the sufficiency of (22) later. The necessity can be shown as follows. If π is an odd prime dividing a_0 the Legendre symbol $(\pi/p) = -1$ by Theorem 5. This is equivalent to $(-p/\pi) = -1$ by the quadratic reciprocity law and $p \equiv 3 \pmod 4$. From this it follows that the highest power of π which divides any value of the quadratic form $px^2 + y^2$, x and y in $\mathfrak{g}$, is even. Multiply (18) by $4g^2$, apply (16) and complete the square. There results

$$p\{(2h - c)k_1 - 2g^2 k_2\}^2 + c^2 k_1^2 = 4g^2(a_0 - k_3 c^2 p). \tag{23}$$

If π divides c, evidently the highest power of π which divides (23) is odd since π^2 does not divide a_0. This contradiction, with a similar one in case 2 divides a_0, implies (22).

Henceforth let $c \geqq 1$ be fixed and assume (22). To determine all $\mathfrak{M}$ whose intersection with P is $\mathfrak{m}_c$ we have first, by Theorem 6, to determine all $\mathfrak{n} = \mathfrak{n}(g, h)$ of the form (16) and (21) such that (18) has solutions and then, for a fixed $\mathfrak{n}$, to determine the effect of taking distinct solutions of (18). We now seek a criterion for the $\mathfrak{n}(g, h)$ such that (18) has solutions.

* Schilling, Mathematische Annalen, vol. 111 (1935), p. 376.

The quadratic form

$$f = f(k_1, k_2) = g_3k_1^2 - p(2h - c)k_1k_2 + pg^2k_2^2 \tag{24}$$

belongs to a class* of quadratic forms of discriminant $-c^2p$, $p \equiv 3 \pmod 4$, uniquely determined by $\mathfrak{n}$. For, $f = c^2pN(k_2\gamma - k_1\nu)$ and a unimodular substitution on the $\mathfrak{g}$-basis (21) of $\mathfrak{n}$ corresponds to a similar substitution on k_1 and k_2. A prime which divides each of g_3, $p(2h-c)$ and pg^2 divides c by (16). Hence, by (22), *a necessary condition that* (18) *have solutions is that f be primitive*. Then f has the following characters.† If $c>1$, let

$$c = 2^\alpha p^\beta q_1^{\alpha_1} \cdots q_s^{\alpha_s}, \qquad (q_1 \cdots q_s, 2p) = 1, \tag{25}$$

be the canonical factorization of c into prime powers. If $\alpha = 0$ or 1, f has the (only) characters (f/p) and (f/q_i) $(i = 1, \cdots, s)$. If $\alpha = 2$, f has the (only) additional character $\delta(n) = (-1)^{(n-1)/2}$, n odd and represented by f. If $\alpha \geqq 3$, f has the (only) additional characters δ and $\epsilon(n) = (-1)^{(n^2-1)/8}$, n odd and represented by f. Since $(a_0, c^2p) = 1$ by (22) and Theorem 5, the Legendre symbols (a_0/p) and (a_0/q_i) and the quantities $\delta(a_0)$, $\epsilon(a_0)$ if δ and ϵ occur for f, define a total character $C(a_0)$ for the discriminant $-c^2p$. The congruence $f \equiv a_0 \pmod{c^2p}$ is easily shown to have solutions if and only if the characters of f have the values prescribed by $C(a_0)$. This is the criterion sought. We have proved

THEOREM 8. *Let c be fixed and satisfy* (22). *Then* (18) *has solutions for all and only those* $\mathfrak{m}_c$*-moduls* $\mathfrak{n}(g, h)$ *for which the associated forms f are primitive and such that their characters have the values prescribed by* $C(a_0)$.

The conditions of Theorem 8 lead to necessary conditions on the integers g. We first prove the

LEMMA. *The form f associated with* $\mathfrak{n}(g, h)$ *is primitive if and only if h can be chosen so that*

$$(g_3, c^2p) = 1. \tag{26}$$

If (26) *holds* (18) *has solutions if and only if* $C(g_3) = C(a_0)$.

The sufficiency of (26) for the primitivity of f is obvious by an earlier remark. To prove the necessity, first let $\beta \geqq 0$. Then (18) requires $(g_3, p) = 1$ and (16) implies $g = p^\beta g_0$, $(g_0, p) = 1$, $h = p^\beta h_0$. We cancel $p^{2\beta}$ in (16) and have $(g_3, p) = 1$. Moreover, $(g_3/p) = 1$. Second, suppose q_i, for a fixed i, does not

* If $\mathfrak{n}$ is a regular $\mathfrak{m}_c$-modul, that is, an $\mathfrak{m}_c$-ideal, the class to which f belongs is primitive and corresponds in the usual way (see Fricke, loc. cit.) to the ideal-class of the order $\mathfrak{m}_c$ to which $\mathfrak{n}$ belongs.

† See Dickson, *Introduction to the Theory of Numbers*, pp. 82–87, or, *Studies in the Theory of Numbers*, p. 37.

divide g. Then, without altering $\mathfrak{n}(g, h)$, we can take $(h, q_i)=1$ and have $(g_3, q_i)=1$ trivially. Moreover, then $(g_3/q_i)=(p/q_i)$. If q_i divides g, (16) implies q_i divides h. Then the primitivity of f requires $(g_3, q_i)=1$. Since h is uniquely determined modulo g^2 by $\mathfrak{n}$, in this case g_3 is uniquely determined modulo q_i by $\mathfrak{n}$. Third, if $\alpha>0$, we proceed for factors 2 as for q_i. In this case, if g is odd, we obtain the additional results that $g_3\equiv p \pmod 4$ or $\pmod 8$ in case $\alpha=2$ or $\alpha\geqq 3$, respectively, and if g is even, g_3 is uniquely determined modulo 4 or 8 in case $\alpha=2$ or $\alpha\geqq 3$, respectively.

The last part of the lemma is obvious since g_3 is represented by f and, if (26) holds, $C(g_3)$ is the total character of f.

The additional conditions on g_3 stated in the proof of the lemma, together with the lemma, lead easily to the following theorem, the details of whose proof we omit.

THEOREM 9. *An integer g, such that* (16) *has solutions, leads to $\mathfrak{m}_c$-moduls $\mathfrak{n}(g, h)$ satisfying the conditions of Theorem* 8 *only if*

$$g = 2^{\alpha_0} p^{\beta} Q_1 g_0, \qquad (g_0, p) = 1, \tag{27}$$

where β is given in (25), *Q_1 is the product of the prime powers $q_i^{\alpha_i}$ in* (25) *for which $(p/q_i)\neq(a_0/q_i)$, and $\alpha_0=0$ if $\alpha=0$ or 1, $\alpha_0=0$ or 1 if $\alpha=2$ according as $p-a_0\equiv 0$ or 2 (mod 4) and $\alpha_0=0$, $\alpha-1$, $\alpha-2$, or $\alpha-1$ if $\alpha\geqq 3$ according as $p-a_0\equiv 0$, 2, 4, or 6 (mod 8), respectively.*

It should be noted that g_0 is required by Theorem 9 to be prime to p, but not necessarily prime to $2Q_1$. Naturally, g_0 must be such that (16) have a solution h. The sufficiency of (22) in Theorem 7 can be proved by showing that there exist $\mathfrak{n}(g, h)$ satisfying Theorem 8 for g given by (27) with $g_0=1$. We omit the details of this proof.

We now assume that $\mathfrak{n}(g, h)$ is fixed and such that (18) has solutions. Let $\lambda^{(1)}$ and $\lambda^{(2)}$ be given by (17) with solutions $k_1^{(1)}$, $k_2^{(1)}$ and $k_1^{(2)}$, $k_2^{(2)}$, respectively, of (18). It is plain that $\mathfrak{M}(\lambda^{(1)})=\mathfrak{M}(\lambda^{(2)})$ if and only if

$$(\lambda^{(2)} - u)\mathfrak{n} \subset \mathfrak{M}(\lambda^{(1)}) \tag{28}$$

since $\mathfrak{M}(\lambda^{(2)})$ is maximal. It is readily shown that (28) holds if and only if $r(\lambda^{(2)}-\lambda^{(1)})$ and $rv(\lambda^{(2)}-\lambda^{(1)})$ are in $\mathfrak{m}_c$ and that the conditions, written in (29) below, on the coefficients of ω in these expressions are necessary and sufficient. In this way we obtain

THEOREM 10. *Let c and $\mathfrak{n}(g, h)$ be fixed and such that* (18) *has solutions. Then for two solutions $k_1^{(1)}$, $k_2^{(1)}$ and $k_1^{(2)}$, $k_2^{(2)}$ of* (18), *and $\lambda^{(1)}$ and $\lambda^{(2)}$ defined by* (17) *with these solutions, $\mathfrak{M}(\lambda^{(1)})=\mathfrak{M}(\lambda^{(2)})$ if and only if, simultaneously,*

$$(29)\qquad \begin{aligned}(2h-c)(k_1^{(2)}-k_1^{(1)})-2g^2(k_2^{(2)}-k_2^{(1)})&\equiv 0 \pmod{c^2},\\ 2g_3(k_1^{(2)}-k_1^{(1)})-p(2h-c)(k_2^{(2)}-k_2^{(1)})&\equiv 0 \pmod{c^2p}.\end{aligned}$$

The conditions found in this section become very simple when $c=1$. Then, by Theorem 9, we must have $(g, p)=1$. For every $\pi(g, h)$ with $(g, p)=1$ and $c=1$, the form f has the (only) character $(f/p)=1=(a_0/p)$, and (18) always has exactly two distinct solutions modulo p which do not satisfy (29). Moreover, without altering (18) modulo p, we may take $k_1\equiv 0 \pmod{2g^2}$ and then take k_2 such that $e_1=0$. Then λ is in $\mathfrak{g}$. The special maximal orders found by Albert (loc. cit.) are the $\mathfrak{M}$ of Theorem 6 with $c=g=1$. Except for certain particular algebras Q (cf. Schilling, loc. cit.) it is not known that an arbitrary maximal order $\mathfrak{M}$ of Q contains the maximal order of some canonical splitting field P, nor, indeed, that $\mathfrak{M}$ contains the maximal order of any splitting field of Q.

UNIVERSITY OF CHICAGO,
CHICAGO, ILL.

ON THE ZEROS OF SUCCESSIVE DERIVATIVES OF INTEGRAL FUNCTIONS*

BY

I. J. SCHOENBERG

INTRODUCTION

1. Let $\eta_0, \eta_1, \eta_2, \cdots$ be a sequence of complex numbers and $f(x)$ an analytic function. The expansion in series of polynomials

$$f(x) = \sum_{n=0}^{\infty} f^{(n)}(\eta_n) p_n(x), \tag{1}$$

where

$$p_n(x) = \int_{\eta_0}^{x} dx_1 \int_{\eta_1}^{x_1} dx_2 \cdots \int_{\eta_{n-1}}^{x_{n-1}} dx_n \qquad (n \geqq 1,\ p_0(x) = 1), \tag{2}$$

was made the subject of important recent investigations by W. Goncharoff† and J. M. Whittaker.‡ Goncharoff calls (1) a generalized Abel series; it is readily found to be valid if $f(x)$ is a polynomial and reduces to the Taylor expansion of $f(x)$ about the point $x=a$ if $\eta_0=\eta_1=\cdots=a$. Whittaker investigated the expansion (1) in the case when every $\eta_n=1$ or -1§ and proved the following theorem:‖

Let $f(x)$ be an integral function for which

$$\overline{\lim_{r\to\infty}} \frac{\log M(r)}{r} = \gamma < \frac{1}{2}, \tag{3}$$

$M(r)$ denoting the maximum modulus of the function on the circle $|x|=r$. Let $\{\eta_n\}$ be a sequence with $\eta_n=\pm 1$. Then the expansion (1) is valid in the circle

---

* Presented to the Society, January 2, 1936; received by the editors November 19, 1935.

† W. Goncharoff, *Recherches sur les dérivées successives des fonctions analytiques. Généralisation de la série d'Abel*, Annales Scientifiques de l'École Normale, vol. 47 (1930), pp. 1–78.

‡ J. M. Whittaker, *On Lidstone's series and two-point expansions of analytic functions*, Proceedings of the London Mathematical Society, (2), vol. 36 (1933–34), pp. 451–469.

§ Whittaker actually considers the case $\eta_n=0$ or 1, which may be reduced to the case $\eta_n=\pm 1$ by an obvious linear transformation of the x-plane.

‖ Loc. cit., Theorem 7, p. 468. A weaker form of this theorem, the number 1/2 on the right side of (3) being replaced by $1/(2e)$, is a particular case of a general theorem proved by Goncharoff, loc. cit., pp. 35–36, and by S. Takenaga, Proceedings of the Physico-Mathematical Society of Japan, vol. 13 (1931), pp. 111–132.

$$|x| < \frac{1}{\gamma} - 1. \tag{4}$$

An immediate corollary stated by Whittaker is as follows.

If $f(x)$ is an integral function satisfying the condition (3) *and*

$$f^{(n)}(-1)f^{(n)}(1) = 0 \qquad (n = 0, 1, 2, \cdots), \tag{5}$$

then $f(x) = 0$ identically.

2. Whittaker remarks, by consideration of the function $\sin(x+1)\pi/4$, of type $\gamma = \pi/4$, which satisfies (5) without vanishing identically, that the number $\frac{1}{2}$, on the right side of (3), can not be replaced by any number larger than $\pi/4$. He conjectures that $\pi/4$ is the best constant in the condition (3). A verification of this conjecture is contained in the following

THEOREM 1. *Let $f(x)$ be an integral function for which*

$$\overline{\lim_{r\to\infty}} \frac{\log M(r)}{r} < \frac{\pi}{4}. \tag{6}$$

Let $\{\eta_n\}$ be a sequence of real numbers, with $-1 \leqq \eta_n \leqq 1$. Then the expansion (1) *is valid throughout the complex plane, absolutely and uniformly in every finite region.*

As above we have the following

COROLLARY. *If an integral function $f(x)$, satisfying* (6), *is such that every one of its derivatives $f(x), f'(x), f''(x), \cdots$ vanishes somewhere between -1 and $+1$, then $f(x) = 0$ identically.*

It might be of interest to point out the following consequence which is obtained by transforming these results by means of $z = (x+1)/4$:

Among all integral functions $f(z) \not\equiv 0$ which are such that $f(z), f'(z), f''(z), \cdots$ vanish somewhere within the interval $0 \leqq z \leqq \frac{1}{2}$, the function $\sin \pi z$ *is the one of least growth.*

3. To establish Theorem 1 we shall use essentially the same tools with which Whittaker proved his results stated above. However, these did not suffice entirely in the present problem for the following reason. Whittaker could solve completely by his method the questions regarding the validity of Lidstone's series (loc. cit., Theorem 2, p. 455) and of (1) in the case when $\eta_n = (-1)^n$ (loc. cit., §6, p. 458), because the polynomials of both these expansions reduce essentially to Appell sequences, and could thus be conveniently estimated by means of their generating functions. Our polynomials (1) $(-1 \leqq \eta_\nu \leqq 1)$ do not form an Appell sequence, for, in the successive integra-

tions in (2), the lower limits of integration proceed in the wrong direction, so to speak. This difficulty is overcome by the consideration of the polynomials (21) which do form an Appell sequence. The solution of an extremal problem (Lemma 1) leads to an estimate for the polynomials (21) (Lemma 2) which holds uniformly in the η's within $(-1, 1)$; then surely this result may be applied to the polynomials (2) as well, provided $-1 \leqq \eta_\nu \leqq 1$.

I. An extremal problem concerning the polynomials $p_n(x)$

4. An important step in the proof of Theorem 1 is the solution of the following extremal problem: *Let $p_n(x)$ be a real polynomial of degree $n(>0)$ of the form $p_n(x)=x^n/n!+\cdots$. Let $p_n(x)$, $p_n'(x), \cdots, p_n^{(n-1)}(x)$ all have at least one zero within the interval $-1 \leqq x \leqq 1$. What is the largest value of $|p_n(0)|$?*

This problem is solved by the following

Lemma 1. *The largest possible value of $|p_n(0)|$ is $|P_n(0)|$, where*

$$P_n(x) = \int_{-1}^{x} dx_1 \int_{+1}^{x_1} dx_2 \cdots \int_{(-1)^n}^{x_{n-1}} dx_n. \tag{7}$$

As

$$p_n(x) = \int_{\eta_0}^{x} dx_1 \int_{\eta_1}^{x_1} dx_2 \cdots \int_{\eta_{n-1}}^{x_{n-1}} dx_n \qquad (-1 \leqq \eta_\nu \leqq 1) \tag{8}$$

is the general type of polynomials considered in our extremal problem, Lemma 1 is equivalent with the following inequality

$$\left| \int_{-1}^{0} dx_1 \int_{+1}^{x_1} dx_2 \cdots \int_{(-1)^n}^{x_{n-1}} dx_n \right| \geqq \left| \int_{\eta_0}^{0} dx_1 \int_{\eta_1}^{x_1} dx_2 \cdots \int_{\eta_{n-1}}^{x_{n-1}} dx_n \right|, \tag{9}$$

$$\text{if } -1 \leqq \eta_\nu \leqq 1.$$

We shall find it more convenient, however, to prove the more inclusive inequality

$$\left| \int_{-1}^{x} dx_1 \int_{+1}^{x_1} dx_2 \cdots \int_{(-1)^n}^{x_{n-1}} dx_n \right| \geqq \left| \int_{\eta_0}^{x} dx_1 \int_{\eta_1}^{x_1} dx_2 \cdots \int_{\eta_{n-1}}^{x_{n-1}} dx_n \right|, \tag{10}$$

$$\text{if } 0 \leqq x \leqq 1, \ -1 \leqq \eta_\nu \leqq 1.$$

We shall therefore prove that *for any fixed x, with $0 \leqq x \leqq 1$, $|P_n(x)|$ is the largest possible value of $|p_n(x)|$.**

Let

* This implies, for reasons of symmetry, that, if $-1 \leqq x \leqq 0$, the largest possible value of $|p_n(x)|$ is $|P_n(-x)|$. The answer to our extremal problem is an entirely different one if x has a fixed value, with $|x| > 1$, and is of no interest in the present investigation.

$$\Phi_n(x;\eta) = \Phi_n(x;\eta_0,\eta_1,\cdots,\eta_{n-1}) \tag{11}$$
$$= \left|\int_{\eta_0}^{x} dx_1 \left|\int_{\eta_1}^{x_1} dx_2 \cdots \left|\int_{\eta_{n-1}}^{x_{n-1}} dx_n\right| \cdots \right|\right|.$$

From (7), (8), and (11) we get

$$\Phi_n(x;\eta) \geqq |p_n(x)|, \qquad \Phi_n(x;-1,\cdots,(-1)^n) = |P_n(x)|;$$

hence (10) will be proved if we can show that

$$\Phi_n(x;-1,\cdots,(-1)^n) \geqq \Phi_n(x;\eta_0,\cdots,\eta_{n-1}), \tag{12}$$
$$\text{for } 0 \leqq x \leqq 1,\ -1 \leqq \eta_\nu \leqq 1.$$

Let $x (0 \leqq x \leqq 1)$ be kept fixed; $\Phi_n(x;\eta)$ is a continuous function of $(\eta_0,\cdots,\eta_{n-1})$ and (12) states that it reaches its maximum value within the cube $-1 \leqq \eta_\nu \leqq 1$ at a certain corner, namely, $(\eta) = (-1,\cdots,(-1)^n)$. We shall prove now that *the maximum value within the cube is necessarily attained at some one of the 2^n corners $(\eta_\nu = \pm 1)$ of the cube.* For let $(\eta_0,\eta_1,\cdots,\eta_{n-1})$ be a point at which the maximum within the cube is reached. From (11) it is obvious that $-1 < \eta_0 < 1$ is impossible, for $\Phi_n(x;\eta)$ can be further increased by increasing the difference $|x-\eta_0|$. Hence $\eta_0 = \pm 1$. Let

$$\eta_\nu = \pm 1 \qquad (\nu = 0, 1, \cdots, k-1), \qquad -1 < \eta_k < 1, \tag{13}$$

where k may take any value from 1 to $n-1$. In view of (13) we can write

$$\Phi_n(x;\eta) = \int\cdots\int_{\mathcal{D}} \left|\int_{\eta_k}^{x_k} \Phi_{n-k-1}(x_{k+1};\eta_{k+1},\cdots,\eta_{n-1})dx_{k+1}\right| dx_1\cdots dx_k, \tag{14}$$

where the integrand $\left|\int_{\eta_k}^{x_k}\Phi_{n-k-1}dx_{k+1}\right|$ is integrated in the ordinary sense over a certain polyhedral domain $\mathcal{D}$, in the space of the variables $(x_1,\cdots,x_k)$, which is entirely contained within the cube $-1 \leqq x_\nu \leqq 1$ $(\nu = 1, 2, \cdots, k)$.* In equation (14) we now keep everything fixed but η_k, which may vary from -1 to $+1$. In view of the formula

$$\frac{\partial}{\partial \eta_k}\left|\int_{\eta_k}^{x_k} \Phi_{n-k-1}(x_{k+1};\eta_{k+1},\cdots,\eta_{n-1})dx_{k+1}\right|$$
$$= \Phi_{n-k-1}(\eta_k;\eta_{k+1},\cdots,\eta_{n-1})\ \operatorname{sgn}(\eta_k - x_k),$$

(14) gives

---

* Thus if $k=3$, $\eta_0=-1$, $\eta_1=-1$, $\eta_2=+1$, then

$$\Phi_n(x;\eta) = \int_{-1}^{x} dx_1 \int_{-1}^{x_1} dx_2 \int_{x_2}^{1} \left|\int_{\eta_3}^{x_3} \Phi_{n-4} dx_4\right| dx_3$$

and the domain $\mathcal{D}$ is defined by the inequalities

$$-1 \leqq x_1 \leqq x, \qquad -1 \leqq x_2 \leqq x_1, \qquad x_2 \leqq x_3 \leqq 1.$$

$$\frac{\partial}{\partial \eta_k}\Phi_n(x;\eta) = \Phi_{n-k-1}(\eta_k;\eta_{k+1},\cdots,\eta_{n-1})\int\cdots\int_{\mathcal{D}}\operatorname{sgn}(\eta_k - x_k)dx_1\cdots dx_k. \tag{15}$$

The integral on the right side of (15) is a non-decreasing continuous function of η_k which has the value ($-$ volume $\mathcal{D}$) for $\eta_k=-1$ and ($+$ volume $\mathcal{D}$) for $\eta_k=1$; moreover this non-decreasing function may have only one interval of constancy at one end of the interval $(-1, 1)$ and the function can therefore vanish at one point only, say $\eta_k=\xi$. The first factor $\Phi_{n-k-1}(\eta_k;\cdots)$ on the right side of (15) is a continuous and positive function of η_k which vanishes only for $\eta_k=\eta_{k+1}$. Hence $\partial\Phi_n/\partial\eta_k$, as given by (15), vanishes for $\eta_k=\xi$, $\eta_k=\eta_{k+1}$; for any other value of η_k, however, it is <0 or >0, according as $\eta_k<\xi$ or $\eta_k>\xi$. Hence, as η_k varies from -1 to $+1$, $\Phi_n(x;\eta)$ will decrease to a certain minimum value to increase thereafter for the remainder of this interval. This shows that our assumption (13), namely, $-1<\eta_k<1$, is impossible. *The point where $\Phi_n(x;\eta)$ reaches its maximum must therefore be found among the 2^n corners of the cube $-1\leqq\eta_\nu\leqq 1$.*

5. It suffices therefore to prove the inequality (12) under the restricted assumptions that all $\eta_\nu=\pm 1$ (and $0\leqq x\leqq 1$, as throughout our discussion). Now all but one set of the absolute value signs in (11) may be dropped and we can write

$$\Phi_n(x;\eta) = \left|\int_{\eta_0}^{x} dx_1\int_{\eta_1}^{x_1} dx_2\cdots\int_{\eta_{n-1}}^{x_{n-1}} dx_n\right| \qquad (\eta_\nu = \pm 1).$$

Since

$$\begin{aligned}\Phi_n(x;1,\eta_1,\cdots,\eta_{n-1}) &= \Phi_n(-x;-1,-\eta_1,\cdots,-\eta_{n-1})\\ &\leqq \Phi_n(x;-1,-\eta_1,\cdots,-\eta_{n-1}),\end{aligned}$$

it suffices to consider the case $\eta_0=-1$ only. For $n=1$ (12) is true. *We assume now* (12) *to hold (for $\eta_\nu=\pm 1$) if n is replaced by any number $m<n$.* We shall prove the inequalities

$$\Phi_n(x;-1,+1,\cdots,(-1)^n) \geqq \Phi_n(x;-1,+1,\eta_2,\cdots,\eta_{n-1}), \qquad (\eta_\nu=\pm 1), \tag{16}$$

$$\Phi_n(x;-1,+1,\cdots,(-1)^n) \geqq \Phi_n(x;-1,-1,\eta_2,\cdots,\eta_{n-1}), \qquad (\eta_\nu=\pm 1), \tag{17}$$

which cover all remaining possibilities.

To prove (16) we write

$$\Phi_n(x; -1, \cdots, (-1)^n) = \int_{-1}^{0} \Phi_{n-1}(x_1; 1, -1, \cdots)dx_1 + \int_{0}^{x} \Phi_{n-1}(x_1; 1, -1, \cdots)dx_1,$$

or

$$\Phi_n(x; -1, \cdots, (-1)^n) = \int_{-1}^{0} \Phi_{n-1}(-x_1; -1, +1, \cdots)dx_1 + \int_{0}^{x} dx_1 \int_{x_1}^{1} \Phi_{n-2}(x_2; -1, +1, \cdots)dx_2, \tag{18}$$

and correspondingly

$$\Phi_n(x; -1, +1, \eta_2, \cdots, \eta_{n-1}) = \int_{-1}^{0} \Phi_{n-1}(-x_1; -1, -\eta_2, \cdots)dx_1 + \int_{0}^{x} dx_1 \int_{x_1}^{1} \Phi_{n-2}(x_2; \eta_2, \cdots, \eta_{n-1})dx_2. \tag{19}$$

In the first integrals in (18) and (19) we have $0 \leqq -x_1 \leqq 1$, and therefore the integrand in (18) is $\geqq$ the integrand in (19), on account of our assumption. In the second integrals in (18) and (19) we have $0 \leqq x_2 \leqq 1$ and for a similar reason the proof of (16) is completed.

In order to prove (17) we consider the function

$$g(x) = \Phi_n(x; -1, +1, \cdots, (-1)^n) - \Phi_n(x; -1, -1, \eta_2, \cdots, \eta_{n-1})$$

and we have to prove that

$$g(x) \geqq 0 \quad \text{if} \quad 0 \leqq x \leqq 1. \tag{20}$$

This will be done in three steps. We show first that $g'(x)$ is a non-increasing function. Indeed

$$g(x) = \int_{-1}^{x} dx_1 \int_{x_1}^{1} \Phi_{n-2}(x_2; -1, \cdots)dx_2 - \int_{-1}^{x} dx_1 \int_{-1}^{x_1} \Phi_{n-2}(x_2; \eta_2, \cdots)dx_2,$$

whence

$$g'(x) = \int_{x}^{1} \Phi_{n-2}(x_2; -1, \cdots)dx_2 - \int_{-1}^{x} \Phi_{n-2}(x_2; \eta_2, \cdots)dx_2,$$

which proves the statement, for $g'(x)$ is a sum of two non-increasing functions. In view of this fact, (20) will be established if we can show that $g(0) \geqq 0$ and $g(1) \geqq 0$, i.e., that (17) holds for $x=0$ and $x=1$. To prove (17) for $x=0$ we notice that

$$\Phi_n(0; -1, 1, \cdots, (-1)^n) = \int_{-1}^{0} \Phi_{n-1}(x_1; 1, -1, \cdots)dx_1$$
$$= \int_{-1}^{0} \Phi_{n-1}(-x_1; -1, +1, -1, \cdots)dx_1;$$

similarly

$$\Phi_n(0; -1, -1, \eta_2, \cdots) = \int_{-1}^{0} \Phi_{n-1}(-x_1; 1, -\eta_2, \cdots)dx_1,$$

and the integrand of the second integral is superior ($\geqq$) to the integrand of the third integral since $0 \leqq -x_1 \leqq 1$. To prove (17) for $x=1$ we notice that

$$\Phi_n(1; -1, -1, \eta_2, \cdots) = \int_{-1}^{1} \Phi_{n-1}(x_1; -1, \eta_2, \cdots)dx_1$$
$$= \int_{-1}^{1} \Phi_{n-1}(-x_1; 1, -\eta_2, \cdots)dx_1$$
$$= \int_{-1}^{1} \Phi_{n-1}(x_1; 1, -\eta_2, \cdots)dx_1$$
$$= \Phi_n(1; -1, 1, -\eta_2, \cdots)$$
$$\leqq \Phi_n(1; -1, 1, -1, \cdots, (-1)^n),$$

the last inequality resulting from (16) for $x=1$. This completes the proof of Lemma 1.

II. An estimate of $p_n(x)$ of lowest order in n for $n \to \infty$

6. Instead of the polynomials $p_n(x)$ defined by (2) we shall now investigate the sequence of polynomials

$$(21) \quad q_n(x) = \int_{\eta_{n-1}}^{x} dx_1 \int_{\eta_{n-2}}^{x_1} dx_2 \cdots \int_{\eta_0}^{x_{n-1}} dx_n \qquad (-1 \leqq \eta_\nu \leqq 1; q_0(x) = 1).$$

This is a sequence of Appell polynomials, i.e., satisfying the recurrence relation $q_n'(x) = q_{n-1}(x)$. It is known and easily shown* that the $q_n(x)$ have a generating function of the type

$$(22) \qquad e^{xt}f(t) = \sum_{n=0}^{\infty} q_n(x)t^n,$$

where

* P. Appell, *Sur une classe de polynômes*, Annales de l'École Normale, (2), vol. 9 (1880), pp. 119–144.

$$f(t) = 1 + \frac{a_1}{1!}t + \frac{a_2}{2!}t^2 + \cdots . \tag{23}$$

This series, if multiplied formally with the power series expansion of e^{xt}, gives

$$q_n(x) = \frac{1}{n!}\left\{x^n + \binom{n}{1}a_1x^{n-1} + \cdots + a_n\right\}. \tag{24}$$

A particular sequence of polynomials of the form (21) will be of special interest, namely,

$$r_n(x) = \int_{(-1)^n}^{x} dx_1 \int_{(-1)^{n-1}}^{x_1} dx_2 \cdots \int_{-1}^{x_{n-1}} dx_n \qquad (r_0(x) = 1). \tag{25}$$

To find the generating function

$$e^{xt}g(t) = \sum_{n=0}^{\infty} r_n(x)t^n \tag{26}$$

of the polynomials $r_n(x)$ we note that they are characterized among all Appell sequences by the conditions

$$r_n(+1) = 0 \quad (n = 2, 4, 6, \cdots), \qquad r_n(-1) = 0 \quad (n = 1, 3, 5, \cdots);$$

these conditions are equivalent to the requirement that $e^tg(t) - 1$ be an odd function of t and $e^{-t}g(t)$ be an even function, i.e.,

$$e^tg(t) - 1 = -e^{-t}g(-t) + 1, \qquad e^{-t}g(t) = e^tg(-t),$$

whence we get

$$g(t) = e^t \operatorname{sech} 2t.$$

Now (26) becomes

$$e^{(x+1)t} \operatorname{sech} 2t = \sum_{n=0}^{\infty} r_n(x)t^n. \tag{27}$$

7. A convenient expression for $r_n(x)$ can now be found by a customary method. Let

$$r_n(x;\rho) = \frac{1}{2\pi i}\int_{|t|=\rho} t^{-n-1}e^{(x+1)t} \operatorname{sech} 2t\, dt, \tag{28}$$

where the integration is taken in the positive sense of rotation along the circumference $|t| = \rho$.

Noting that $t = \pm\pi i/4$ are the only (simple) poles of sech $2t$ within the circle $|t| \leqq 1$, we find by Cauchy's theorem of residues

$$r_n(x;1) - r_n\left(x;\frac{\pi}{8}\right) = \frac{(-1)^{[n/2]+1}}{2}\left(\frac{4}{\pi}\right)^{n+1}\begin{Bmatrix}\cos\\ \sin\end{Bmatrix}\frac{\pi}{4}(x+1), \quad \begin{Bmatrix}n \text{ even}\\ n \text{ odd}\end{Bmatrix},$$

where on the right side we have the factor $\cos (x+1)\pi/4$ or $\sin (x+1)\pi/4$, according as n is even or odd. In view of (27) we have

$$r_n\left(x;\frac{\pi}{8}\right) = r_n(x)$$

and finally

$$(29)\quad r_n(x) = \frac{(-1^{[n/2]})}{2}\left(\frac{4}{\pi}\right)^{n+1}\begin{Bmatrix}\cos\\ \sin\end{Bmatrix}\frac{\pi}{4}(x+1) + r_n(x;1), \quad \begin{Bmatrix}n \text{ even}\\ n \text{ odd}\end{Bmatrix},$$

where

$$(30)\qquad |r_n(x;1)| \leqq (2\pi)^{-1}e^{|x|}\int_{|t|=1} |e^t \operatorname{sech} 2t| \cdot |dt|.$$

Formulas (29) and (30) reveal some interesting properties of these polynomials, as, for example, the relations

$$(31)\quad \lim_{n\to\infty}\frac{r_{2n}(x)}{r_{2n}(-1)} = \cos\frac{\pi}{4}(x+1), \qquad \lim_{n\to\infty}\frac{r_{2n+1}(x)}{r_{2n+1}(1)} = \sin\frac{\pi}{4}(x+1),$$

which hold uniformly throughout any finite domain. For our present purpose we shall need only an estimate for $|r_n(x)|$. Formula (29) is equivalent to

$$r_n(x) = (-1)^{[n/2]}\left(\frac{4}{\pi}\right)^{n+1}2^{-3/2}\left(\cos\frac{\pi x}{4} + (-1)^{n-1}\sin\frac{\pi x}{4}\right) + r_n(x;1);$$

as both $|\cos(\pi x/4)|$ and $|\sin(\pi x/4)|$ are $\leqq \cosh(\pi|x|/4)$, it follows that

$$(32)\qquad |r_n(x)| < \left(\frac{4}{\pi}\right)^{n+1}2^{-1/2}\cosh\frac{\pi|x|}{4} + Ke^{|x|},$$

where

$$(33)\qquad K = \frac{1}{2\pi}\int_{|t|=1} |e^t \operatorname{sech} 2t| \cdot |dt|.$$

8. From (32) an estimate for $q_n(x)$, defined by (21), may readily be found. For let

$$\gamma_n = n!\,|r_n(0)|;$$

(25) shows that the sign of $r_n(0)$ is $(-1)^{[n/2]}$, hence the general formula (24) gives

$$r_n(x) =$$

$$\frac{1}{n!}\left\{x^n + \binom{n}{1}\gamma_1 x^{n-1} - \binom{n}{2}\gamma_2 x^{n-2} - \binom{n}{3}\gamma_3 x^{n-3} + \cdots + (-1)^{[n/2]}\gamma_n\right\}.$$

Substituting here $|x|/i$ for x, we get

$$r_n(|x|\, i^{-1})$$

$$= \frac{i^{-n}}{n!}\left\{|x|^n + i\binom{n}{1}\gamma_1 |x|^{n-1} + \binom{n}{2}\gamma_2 |x|^{n-2} + i\binom{n}{3}\gamma_3 |x|^{n-3} + \cdots\right\},$$

whence, by the inequality $|a| + |b| \leqq 2^{1/2}|a+ib|$ for a and b real,

$$\frac{1}{n!}\left\{|x|^n + \binom{n}{1}\gamma_1 |x|^{n-1} + \binom{n}{2}\gamma_2 |x|^{n-2} + \cdots + \gamma_n\right\} \leqq 2^{1/2}|r_n(|x|\, i^{-1})|. \tag{34}$$

By Lemma 1 we have

$$n!\,|q_n(0)| = |a_n| \leqq n!\,|r_n(0)| = \gamma_n,$$

and now (24) and (34) give

$$|q_n(x)| \leqq 2^{1/2}|r_n(|x|\, i^{-1})|.$$

This inequality and (32) give an estimate for $q_n(x)$. As the order of the lower limits of integration η_ν $(-1 \leqq \eta_\nu \leqq 1)$ in (21) is now immaterial, since the result holds for arbitrary η_ν, their order may now be reversed. This proves

LEMMA 2. *If*

$$p_n(x) = \int_{\eta_0}^{x} dx_1 \int_{\eta_1}^{x_1} dx_2 \cdots \int_{\eta_{n-1}}^{x_{n-1}} dx_n \qquad (-1 \leqq \eta_\nu \leqq 1,\ \nu = 0, 1, \cdots, n-1), \tag{35}$$

then

$$|p_n(x)| < \left(\frac{4}{\pi}\right)^{n+1} \cosh\frac{\pi|x|}{4} + 2^{1/2}Ke^{|x|},* \tag{36}$$

where K is the numerical constant defined by (33).

* Goncharoff's general estimate

$$|p_n(x)| \leqq (|x-\eta_0| + |\eta_0-\eta_1| + \cdots + |\eta_{n-2}-\eta_{n-1}|)^n/n!, \tag{36'}$$

which holds for arbitrary values of all quantities involved, is very favorable if η_n tends to a limiting value for $n\to\infty$; it is less favorable, however, if, for example, $\eta_n=(-1)^{n+1}$. To establish his results mentioned in the introduction, Whittaker used the estimate

$$|p_n(x)| \leqq \tfrac{1}{2}e^2(1+|x|)^n \text{ for } |x|>1,\ (\eta_\nu = \pm 1). \tag{36''}$$

This estimate leads only to a finite region of convergence of (1), for functions $f(x)$ with the property (3), because the right side of (36'') is very large for $n\to\infty$, if $|x|$ is large, whereas (36) permits to show that (1) converges in the whole plane if $f(x)$ satisfies (6) and $-1\leqq\eta_\nu\leqq 1$ (§3).

III. Convergence of the Abel series expansion (1) for $-1 \leqq \eta_\nu \leqq 1$

9. Theorem 1 can now easily be proved. We need the following simple and well known lemma.

Lemma 3. *Let $f(x)$ be an integral function for which*

$$\overline{\lim_{r\to\infty}} \frac{\log M(r)}{r} = k < \infty . \tag{37}$$

Then, if $l > k$,

$$|f^{(n)}(0)| < l^n \quad \text{for} \quad n \geqq n_0(l). \tag{38}$$

Take l' such that $k < l' < l$. By Cauchy's inequality

$$|f^{(n)}(0)| \leqq \frac{M(r)n!}{r^n} < \frac{e^{l'r}n!}{r^n} \qquad (r \geqq r_0).$$

Taking $r = n/l'$, this gives

$$|f^{(n)}(0)| < \frac{e^n n!}{n^n} l'^n = \frac{e^n n!}{n^n}\left(\frac{l'}{l}\right)^n \cdot l^n < l^n \qquad (n \geqq n_0).$$

If $f(x)$ is an integral function we have, working formally with the expansion (1),

$$\begin{aligned} \sum_{n=0}^{\infty} f^{(n)}(\eta_n) p_n(x) &= \sum_{n=0}^{\infty}\sum_{r=0}^{\infty} \frac{f^{(n+r)}(0)}{r!} \eta_n^r p_n(x) = \sum_{m=0}^{\infty} f^{(m)}(0) \sum_{n+r=m} \frac{\eta_n^r}{r!} p_n(x) \\ &= \sum_{m=0}^{\infty} f^{(m)}(0) \frac{x^m}{m!} = f(x), \end{aligned} \tag{39}$$

for

$$x^m = \eta_0^m + m\eta_1^{m-1} p_1(x) + \cdots + m! p_m(x) = m! \sum_{n+r=m} \frac{\eta_n^r}{r!} p_n(x),$$

as a result of (1), if applied to $f(x) = x^m$. The formal relations (39) are certainly justified for $|\eta_\nu| \leqq 1$, if the double series

$$\sum_{n=0}^{\infty}\sum_{r=0}^{\infty} \frac{1}{r!} |f^{(n+r)}(0)| \cdot |p_n(x)| \tag{40}$$

is convergent. Let $f(x)$ satisfy (37), then (38) holds for $l > k$. On the other hand (36) gives

$$|p_n(x)| < \left(\frac{4}{\pi}\right)^n A(\rho) + B(\rho) \text{ for } |x| \leqq \rho.$$

Hence (40) is dominated for $|x| \leqq \rho$ by the series

$$\sum_{n=0}^{\infty} \sum_{r=0}^{\infty} \frac{1}{r!} Hl^{n+r} \left\{ \left(\frac{4}{\pi} \right)^n A(\rho) + B(\rho) \right\} = \sum_{n=0}^{\infty} He^{l} l^n \left\{ \left(\frac{4}{\pi} \right)^n A(\rho) + B(\rho) \right\}, \tag{41}$$

H being a suitable constant. If $k < \pi/4$, we can choose l such that $k < l < \pi/4$, hence $4l/\pi < 1$ and also $l < 1$. Then both series (41) and (40) converge and the expansion (1) is valid and absolutely and uniformly convergent in $|x| \leqq \rho$ for ρ arbitrarily large.

Added in proof, June, 1936. The question whether or not the conditions $-1 \leqq \eta_\nu \leqq 1$ of Theorem 1 may be replaced by $|\eta_\nu| \leqq 1$ is still open. A general theorem of S. Takenaka,* if suitably specialized, insures the validity of (1) around the origin under the assumptions $|\eta_\nu| \leqq 1$, provided $\gamma < \log 2$. With the method of the present note the matter hinges upon an extension of Lemma 2 to the case $|\eta_\nu| \leqq 1$. The estimate (36) may well hold under these more general assumptions.

---

* *On the expansion of integral transcendental functions in generalized Taylor's series*, Proceedings of the Physico-Mathematical Society of Japan, (3), vol. 14 (1932), pp. 529–542; see also the recent Cambridge Tract of J. M. Whittaker, *Interpolatory Function Theory*, Cambridge, 1935, where the same problem is pointed out as unsolved on p. 45.

SWARTHMORE COLLEGE,
SWARTHMORE, PA.

PERIODIC ONE-PARAMETER GROUPS IN THREE-SPACE*

BY

DEANE MONTGOMERY† AND LEO ZIPPIN

Introduction. We shall consider a one-parameter group of homeomorphisms of euclidean three-space, here designated by E, into itself. We shall postulate for our group a certain weak periodic character with respect to the points of E and shall show that it is closely related to the rotation group of three-space. In particular it will be easy to determine when it is topologically equivalent to that group. In anticipation of the explanatory remarks of the next section, we may formulate our principal theorem as follows:

A continuous one-parameter group $T(x; t)$,

$$x \epsilon E, \qquad -\infty < t < \infty,$$

of pointwise-periodic homeomorphisms of E *into itself, whose period function is bounded in every sphere, is a topological quasi-rotation group. In particular, it is the topological rotation group whenever the period function is constant over the moving points.*

1. To make clear the meaning of the theorem, let us denote by R a fixed euclidean three-space of reference, given in cylindrical coordinates: (r, θ, z). Let $F(r, z)$ be any positive continuous function defined for all z and all $r>0$. For a fixed real number t, the correspondence

$$T_F(r, \theta, z; t): \qquad (r, \theta, z) \to (r, \theta + 2\pi[F(r, z)]t, z)$$

defines a homeomorphism of R into itself which we shall call a quasi-rotation. The totality of these mappings, for all real t, is a group and we call it a quasi-rotation group. In particular, when $F(r, z)$ is constant, it is the rotation group and is independent, topologically speaking, of the value of the constant. Now let H denote any homeomorphism of R into E. Corresponding to each t, for a fixed $F(r, z)$, H induces a homeomorphism of E into itself: their totality, for all t, will be called a *topological quasi-rotation group* of E. The proof of our theorem will consist in constructing a suitable function F and homeomorphism H.

2. In accordance with our premises, let $T(x; t)$ be a *one-parameter group* of homeomorphisms of E such that

* Presented to the Society, September 12, 1935; received by the editors June 12, 1935.

† National Research Fellow.

(i) $T(x; t)$ is continuous simultaneously in x and t;

(ii) to each solid sphere S in E there corresponds an integer N such that if $x \epsilon S$ there exists at least one $\bar{t} = \bar{t}(x)$ for which

$$T(x; \bar{t}) = x, \qquad 0 < \bar{t} < N.$$

For each x the set of points $T(x; t)$ is a continuous image of the set of real numbers. It will be called the *orbit* of x and designated by O_x. The set of real numbers t for which $T(x; t) = x$ is a closed subgroup (not the identity) of the real additive group, depending on x, and O_x is homeomorphic to the corresponding factor group. Therefore O_x is either a simple closed curve, the homeomorph of a circumference, or the single point x. When $O_x = x$, x will be called a *fixed point* and said to have the period zero. Otherwise x will be called a *moving point* and its period $p(x)$ will be the least positive number for which $T[x; p(x)] = x$. By condition (ii) the period function $p(x)$ is bounded in every sphere. By the group property it is constant over each orbit. The following lemma is quite easy.

LEMMA 1. *If the sequence of points x_n converges to a moving point x, and the sequence of numbers $p(x_n)$ to a number p, then p is a positive integral multiple of $p(x)$.*

As an immediate corollary $p(x)$ is lower semi-continuous.

3. **The orbit-space E^*.** If the point y is in O_x, there is a t, $0 \leqq t < p(x)$, such that $T(x; t) = y$. It is a trivial consequence of the group property that two orbits have no point in common unless they coincide. Further, if any sequence of points x_n converges to a point x, then

$$\overline{\sum O_{x_n}} \supset O_x$$

by condition (i), and

$$\overline{\sum O_{x_n}} - \sum O_{x_n} \subset O_x$$

in virtue of the boundedness of the period function. That is, $O_x = \lim O_{x_n}$. We may conclude that if the points x_n converge to the point at infinity, that is, properly diverge, then any sequence of points drawn from the corresponding orbits must properly diverge. It is convenient to adjoin to E the point at infinity, P_∞, thus obtaining a closed three-sphere $\overline{E}$. We shall extend $T(x; t)$ by the following definition: $T(P_\infty; t) = P_\infty$ for all t. In this way P_∞ becomes its own orbit and is a fixed point. The extended function remains continuous simultaneously and it is still true that if $x_n \to x$, then $O_{x_n} \to O_x$.

We can now construct the important auxiliary space E^* whose points are in 1–1 correspondence with the orbits of $\overline{E}$, and which we may suppose

metrized.† The continuous single-valued correspondence of $\overline{E}$ into E^* we shall denote by L. If a is any point of E^*, the inverse $L^{-1}(a)$ is an orbit, and A is a closed set in E^* when $L^{-1}(A)$ is closed. Finally since E is compact, connected, and locally connected, E^* will have all of these properties. It will be proved later that E^* is *homeomorphic to a closed 2-cell*. We may now define for every a in E^* a function $p^*(a)$ which has as its value the constant value assumed by $p(x)$ on $L^{-1}(a)$. The function $p^*(a)$ is lower semi-continuous.

4. Let F denote the set of fixed points of $\overline{E}$, M the set of moving points, D the subset of M at which $p(x)$ is discontinuous, and C the subset of M where $p(x)$ is continuous. Then $\overline{E}=F+C+D$, these sets being mutually exclusive. It is clear that F is closed, and it is not vacuous since it certainly contains P_∞. We shall assume that F does not exhaust E, as otherwise the analysis of $T(x; t)$ is trivial. Then $M=C+D$ is open and not vacuous. It is clear that $E^*=L(F)+L(C)+L(D)$, the three sets being mutually exclusive.

LEMMA 2. *The set C is open and not vacuous.*

This is an easy consequence of Lemma 1. There is no reason, a priori, why D should be vacuous, but we shall see later that it is.

The system of orbits contained in M is readily shown to be a regular family as defined by H. Whitney.‡ It is in fact a very special type of such families and we shall therefore be free to exploit the "cross sections" which he has introduced to obtain what shall here call a *true section*.

Definition. If A is a subset of E^*, any relatively closed subset K of $L^{-1}(A)$ which *cuts each orbit of $L^{-1}(A)$ exactly once* will be called a true section of $L^{-1}(A)$. It is clear that K will be homeomorphic to A.

By Lemma 2, $D+F$ is closed; it follows that $L(D+F)$ is closed and that its complement $L(C)$ is open in E^*.

LEMMA 3. *If b is any point in $L(C)$ there is a neighborhood U of b such that $L^{-1}(U)$ contains a true section.*

Let U_1 be any neighborhood of b contained in $L(C)$. The orbits filling $L^{-1}(U_1)$ constitute a regular family, so that if q is any point of $L^{-1}(b)$ there exists a set S which is a *cross section through q* of the curves filling $L^{-1}(U_1)$ [in the sense of Whitney]. Let R_n be the closed sphere of center q and radius $1/n$. The set $R_n \cdot S$ is a cross section through q.§ It will now be shown that for some n each orbit contains at most one point of $R_n \cdot S$. Suppose, contrariwise, that in each set $R_n \cdot S$ there is a point q_n such that O_{q_n} cuts $R_n \cdot S$ in a second

† See Hausdorff, *Mengenlehre*, 1927, p. 145, for a definition of distance between closed sets.

‡ Annals of Mathematics, vol. 34 (1933), pp. 244–270.

§ We shall not repeat the definition and properties of the Whitney cross sections, except by implication.

point r_n. There is then for each n a number t_n such that $|t_n| \leqq \frac{1}{2}\, p(q_n)$ and such that $T(q_n; t_n) = r_n$. Since $p(x)$ is bounded in every bounded set, it may be assumed that $\{t_n\}$ converges to a unique limit t_0. The sequences $\{q_n\}$ and $\{r_n\}$ converge to q, so that $q = \lim r_n = \lim T(q_n; t_n) = T(q; t_0)$. On account of the continuity of $p(x)$ and the fact that $|t_n| \leqq \frac{1}{2} p(q_n)$ it follows that $t_0 = 0$. Consider now the arcs $c_n d_n$ made up of points $T(q_n; t)\,[-|t_n| < t < |t_n|]$. These arcs contain q_n and r_n. But if $q'q''$ is any arc of O_q containing q, and λ is any number, it is easy to see that for some n, there is an arc $c_n d_n$ contained in an arc of the λ-neighborhood (in Whitney's sense, loc. cit.) of $q'q''$. Since this is in contradiction with his cross section, we may conclude that there must exist an n_1 such that each orbit which cuts $R_{n_1} \cdot S$ cuts it exactly once.

The set $R_{n_1} \cdot S$ is mapped in a 1–1 manner on the set $L(R_{n_1} \cdot S)$. Because $R_{n_1} \cdot S$ is compact this mapping is a homeomorphism. From the properties of the Whitney section the set $L(R_{n_1} \cdot S)$ must contain all points in some neighborhood U of b. If K is the part of $R_{n_1} \cdot S$ which is mapped into U, K is a true section of $L^{-1}(U)$.

5. In this section we shall investigate the nature of the inverse images of certain simple sets in E^* preparatory to showing that E^* is a closed 2-cell.

LEMMA 4. *If $\alpha\beta$ is an arc in $L(C)$, then $L^{-1}(\alpha\beta)$ contains a true section and is homeomorphic to the product of a circle and a segment.*

About each point b of $\alpha\beta$ choose a neighborhood U such that $L^{-1}(U)$ contains a true section K. Any subarc $\alpha_1\beta_1$ of $\alpha\beta$ which lies in a single U is clearly such that $L^{-1}(\alpha_1\beta_1)$ contains a true section $\alpha_1'\beta_1'$. Since $\alpha\beta$ may be covered by a finite set of the U's mentioned above, there is on $\alpha\beta$ a set of non-overlapping arcs $(\alpha_0\alpha_1)$, $(\alpha_1\alpha_2)$, $\cdots$, $(\alpha_{n-1}\alpha_n)$ whose sum is $\alpha\beta$ and each of which is in a single U. Let $\alpha_{i-1}''\alpha_i'$ be a true section of $L^{-1}(\alpha_{i-1}\alpha_i)$. There is a real number t_1 such that $T(\alpha_1''; t_1) = \alpha_1'$. Consider the arc $\alpha_0''\alpha_1' + T(\alpha_1''\alpha_2'; t_1)$.† This arc is a true section of $L^{-1}(\alpha_0\alpha_1 + \alpha_1\alpha_2)$ obtained by turning the true section of $L^{-1}(\alpha_1\alpha_2)$ so that its first end point matches with the last end point of the true section of $L^{-1}(\alpha_0\alpha_1)$. By making successive turnings of this type, we obtain an arc $\alpha'\beta'$ which is a true section of $L^{-1}(\alpha\beta)$.

If I is the unit interval and C_1 is a circle of circumference 2π we can show that $L^{-1}(\alpha\beta)$ and the product-space $I \times C_1$ are homeomorphic as follows. Let $h(x)$ be any homeomorphic mapping between $\alpha'\beta'$ and I. Suppose now that y is a point of $L^{-1}(\alpha\beta)$ which is on the orbit cutting $\alpha'\beta'$ in x and suppose that $T(x; t) = y$ where $0 \leqq t < p(x)$. Then because of the continuity of $p(x)$ in C, a homeomorphism between $L^{-1}(\alpha\beta)$ and $I \times C_1$ is defined by the correspondence: $y \longleftrightarrow h(x) \times 2\pi t / p(x)$.

† This shall denote the set of all points x such that $x = T(x'; t_1)$ for some $x' \subset \alpha_1''\alpha_2'$.

LEMMA 5. *If the arc $\alpha\beta$ is in $L(C)$ except for the point α which is in $L(F)$, then $L^{-1}(\alpha\beta)$ has a true section and is a closed 2-cell.*

Express $\alpha\beta-\alpha$ as $\sum_{n=1}^{\infty}\sigma_n$ where the σ_n's are consecutive closed arcs, σ_n and σ_{n+1} abutting at a common end point β_n. Let S_n be a true section of $L^{-1}(\sigma_n)$. The end points b_1 and b_1' of S_1 and S_2 respectively, corresponding to β_1, lie on the same orbit. Choose a t^* so that $T(b_1'; t^*)=b_1$. The arc $S_2'=T(x; t^*)$, x ranging over S_2, is also a true section of $L^{-1}(\sigma_2)$, and S_1+S_2' is a true section of $L^{-1}(\sigma_1+\sigma_2)$. Continuing this indefinitely we obtain a true section K of $L^{-1}(\alpha\beta-\alpha)$. It is clear (compare §3) that if f denotes the fixed point $L^{-1}(\alpha)$, $K+f$ is closed and is the desired section. That $L^{-1}(\alpha\beta)$ is a 2-cell is now trivial.

LEMMA 6. *If J is a simple closed curve in $L(C)$, then $L^{-1}(J)$ contains a true section and is a torus.*

We express J as the sum of two arcs $\alpha\gamma\beta$ and $\alpha\gamma'\beta$. Let acb and $a'c'b'$ be true sections of $L^{-1}(\alpha\gamma\beta)$ and $L^{-1}(\alpha\gamma'\beta)$ respectively. Choose t and t' so that

$$T(a; t) = a', \qquad T(b; t') = b'.$$

Let h be a continuous function defined for all x of acb which has the value t at a and t' at b. The set of points

$$T[x; h(x)], \qquad x \text{ in } acb,$$

is a true section of $L^{-1}(\alpha\gamma\beta)$, and is an arc whose end points coincide with a' and b' respectively. Together with the arc $a'c'b'$ it gives the desired true section of $L^{-1}(J)$. That $L^{-1}(J)$ is a torus can now be seen by defining a homeomorphism similar to the one used in Lemma 4.

6. We shall devote this section to proving that D is vacuous. A certain type of set used below will first be described. Let B be the product of a circle and an interval, in other words a bounded cylinder. Let points of the base be specified by an angular coordinate θ $(0\leqq\theta\leqq 2\pi)$ and on one base of the cylinder identify sets of k points of the form $\theta+2n\pi/k$ $(n=0, 1, \cdots, k-1)$. Any homeomorph of this figure is called a twisted strip of index k.

LEMMA 7. *Let $\alpha\beta$ be an arc which, with the exception of α, lies in $L(C)$, and which is such that as x approaches α along $\alpha\beta$, $p^*(x)$ approaches $kp^*(\alpha)$. Then $L^{-1}(\alpha\beta)$ contains a true section and is a twisted strip of index k.*

We observe first that $L^{-1}(\alpha\beta)$ is locally connected. For let g be any point of this set, U_ϵ the open sphere about g of radius ϵ. There is an open arc of the orbit O_g which contains g and lies entirely in U_ϵ. Let $\delta<\epsilon$ be such that the sphere U_δ about g does not intersect O_g except in points of the open arc. Let $\gamma=L(g)$. It is clear, by §3, that there is an arc σ, open in $\alpha\beta$ and containing γ,

such that every orbit in $L^{-1}(\sigma)$ intersects U_δ. Let Q be the subset of $L^{-1}(\sigma)$ consisting of all points x such that $x = T(\bar{x}; \bar{t})$, where $\bar{x}$ is some point of U_δ and $T(\bar{x}; t)$ is in U_ϵ for every t, $0 \leqq t \leqq \bar{t}$ if $t > 0$, and for every t, $0 \geqq t \geqq \bar{t}$ if $t < 0$. It is now easy to see that Q is connected and open in $L^{-1}(\alpha\beta)$. Hence $L^{-1}(\alpha\beta)$ is locally connected.

Therefore if q is any point of $L^{-1}(\alpha)$ and S is a cross section through q of the regular family filling $L^{-1}(\alpha\beta)$, S is locally connected.† Then there must be an arc qr which is contained in S and which has q and no other point on $L^{-1}(\alpha)$.

It will now be shown that qr cuts no orbit more than once. For, if it did, it would contain a proper subarc $r'r$ with the same property. Now $r'r \subset L^{-1}(\alpha'\beta)$, where $\alpha'\beta$ is a proper subarc of $\alpha\beta$. In virtue of Lemma 4, a quite elementary argument will show that there exists a point v of $r'r$ and there exists an arbitrarily small arc σ on the orbit of v, containing v as inner point, such that on arbitrarily near orbits (possibly on O_v itself) there exist arcs arbitrarily near to σ (in the sense of Whitney) which meet $r'r$ (therefore S) in at least two distinct points. Hence the first condition that S be a cross section (Whitney, loc. cit.) fails at the point v. Therefore qr cuts all orbits at most once.

It follows that qr is mapped homeomorphically on a subset $\alpha\beta_1$ of $\alpha\beta$ and $\alpha\beta_1$ is an arc. If β_1 is different from β, there is an arc $\beta_1'\beta'$ which is a true section of $L^{-1}(\beta_1\beta)$. By turning $\beta_1'\beta'$ so that β_1' coincides with r we obtain an arc which we denote by qr' and which is a true section of $L^{-1}(\alpha\beta)$.

That the surface swept out by qr' under the action of the group is a twisted strip of index k may be seen as follows. Let Q be a cylinder which is the product of K and I, K being a circle of circumference 2π and I being a unit interval. Let y be any point of $L^{-1}(\alpha\beta)$ and denote by x' the point in which the orbit through y cuts qr'. Let t be such that $T(x'; t) = y\,[0 \leqq t < p(y)]$. Let $h(x')$ be a homeomorphism carrying qr' into I. Now consider the following correspondence. If y is in $L^{-1}(\alpha\beta - \alpha)$, y corresponds to $h(x') \times 2\pi t/p(y)$. If y is in $L^{-1}(\alpha)$, y corresponds to $h(x') \times 2\pi t/(kp(y))$. This determines a correspondence between $L^{-1}(\alpha\beta)$ and certain points of Q. However, when the proper identifications are made on Q to make it a twisted strip of index k, the above correspondence actually is a homeomorphism of $L^{-1}(\alpha\beta)$ and this twisted strip.

PRINCIPAL LEMMA 1. *The set D is vacuous; that is, $p(x)$ is continuous on the moving points.*

We shall show that $L(D) \subset E^*$ is vacuous. We know that $L(D)$ is closed

† Whitney, loc. cit.

in $L(M)$, hence $L(D)=P+G$ where P is perfect in $L(M)$ and G is countable. There are two cases:

(i) P is vacuous. Here $L(D)=G$ and contains an isolated point b. Since $L(F)$ is closed, there exists in E^* an arc $\beta' b$ which, except for b, belongs to $L(C)$ where $p^*(x)$ is continuous. Since D is, in this case, a one-dimensional F_σ,† it is well known that its complement in E is arcwise connected. Therefore if f^* denotes any point of $L(F)$, there is in E^* an arc $\beta' f^*$ which contains no point of $L(D)$. Let f be the first point of this arc which is in $L(F)$. In the sum of the arcs $b\beta'$ and $\beta' f$ there is an arc $b\beta f$ such that the open arc belongs to $L(C)$.

Now it is quite easy to see that, at an isolated point b in $L(D)$, $p^*(x)$ approaches a unique limit, say $kp^*(b)$, where by Lemma 1, k must be an integer. Then from previous lemmas, $L^{-1}(b\beta)$ is a twisted strip of index k, and $L^{-1}(\beta f)$ is a 2-cell. It is clear that $L^{-1}(b\beta f)$ is homeomorphic to a complex which contains no absolute cycle, but which is a cycle mod k. Such a configuration cannot exist in E, as can be seen from the Alexander Duality Theorem.‡ In this first case we have now arrived at a contradiction.

(ii) P is not vacuous. The function $p^*(a)$ considered relatively to P is lower semi-continuous and there must be a point b in P at which $p^*(a)$ is continuous§ considered relatively to P. Let O be an open set in $L(M)$ which contains b and which is such that if $a\epsilon O$, $p^*(a)<Np^*(b)$ where N is some fixed integer. Assume further that O is so chosen that if $a\epsilon O\cdot P$,

$$|p^*(a)-p^*(b)|<\frac{p^*(b)}{6N}. \tag{1}$$

Let H_i^* be the set of all points a in $L(C)$ such that

$$|ip^*(b)-p^*(a)|<\frac{p^*(b)}{6}. \tag{2}$$

Let $H_i=H_i^*\cdot O$. Assume now that O is chosen so small that it may be written as a sum of (mutually exclusive) sets as follows: $O=P\cdot O+G\cdot O+\sum_1^N H_i$. Because b is in $L(D)$ there must be an integer $k\geqq 2$ such that H_k is not vacuous. Let B denote all those points of the boundary of O_k which lie in O, where O_k is used to denote a component of H_k. Let x_0 be a point in $L^{-1}(O_k)$ and let x_1 and x_2 be two points on distinct orbits in $L^{-1}(O\cdot P)$. There are at least two such points because P is dense in itself. Join x_0 and x_1 by an arc x_0x_1 which is

† By the "Summensatz": Menger, *Dimensionstheorie*, p. 92.

‡ Alexander, these Transactions, vol. 23, p. 333; and Pontrjagin, Annals of Mathematics, vol. 35, p. 904.

§ Hausdorff, *Mengenlehre*, 1927, p. 255.

in $L^{-1}(O)$ and which intersects neither $L^{-1}(G)$ nor x_2 (the set to be avoided is a 1-dimensional F_σ). Join x_0 and x_2 by an arc x_0x_2 which is in $L^{-1}(O)$ and which intersects neither $L^{-1}(G)$ nor the arc previously constructed† (except of course at the point x_0). Let x_3 be the first point of $L^{-1}(P\cdot O)$ on x_0x_1 and let x_4 be the first point of $L^{-1}(P\cdot O)$ on x_0x_2. These points are in $L^{-1}(B)$. The set $L(x_0x_3)+L(x_0x_4)$ is connected and locally connected; let $L(x_3)=a_3$ and $L(x_4)=a_4$. There must be an arc a_3a_4 which is a subset of $L(x_0x_3)+L(x_0x_4)$. Except for a_3 and a_4 this arc lies entirely in O_k. The function $p^*(a)$ considered along the arc a_3a_4 must approach an integral multiple of $p^*(a_3)$, say $k_3p^*(a_3)$, as a approaches a_3 along the arc, and as a approaches a_4 along the arc, $p^*(a)$ must approach an integral multiple of $p^*(a_4)$, say $k_4p^*(a_4)$.

For a in the interior of the arc we must have, since $a\epsilon O_k \subset H_k \subset H_k^*$,

$$|\, kp^*(b) - p^*(a)\,| < \frac{p^*(b)}{6}. \tag{3}$$

Since a_i $(i=3, 4)$ is in $O\cdot P$ and since $k\leqq N$, (1) implies

$$|\, p^*(b) - p^*(a_i)\,| < \frac{p^*(b)}{6k}.$$

This inequality implies

$$|\, kp^*(b) - kp^*(a_i)\,| < \frac{p^*(b)}{6}. \tag{4}$$

From (3) and (4) we have the following:

$$|\, p^*(a) - kp^*(a_i)\,| < \frac{p^*(b)}{3}. \tag{5}$$

For a on the arc, $p^*(a)>p^*(b)$ and therefore (5) implies that k_3 and k_4 both equal k.

Let a_5 be some point on the interior of a_3a_4. The sets $L^{-1}(a_5a_3)$ and $L^{-1}(a_5a_4)$ are both twisted strips of index k. These strips combine to form a set which is a cycle modulo k but which contains no cycle in the absolute sense. Again by the Alexander Duality Theorem such a set cannot be imbedded in $\overline{E}$ and a contradiction has been reached. Therefore both P and G must be vacuous and the set $L(D)$ is vacuous.

7. It will assist us in analyzing F to construct a certain homeomorphism, defined as follows: For any point x, $H(x)$ is the point $T[x;\ p(x)/2]$. That $H(x)$ is a homeomorphism follows from the properties of $T(x;\ t)$ and the fact

† Again the set to be avoided is a 1-dimensional F_σ.

that $p(x)$ is continuous at moving points. If x_0 is a fixed point the continuity of $H(x)$ at x_0 follows even though $p(x)$ is not continuous at x_0 on account of the boundedness required of $p(x)$. The fixed points of $H(x)$ are identical with the fixed points of $T(x; t)$. The transformation $H(x)$ clearly has period 2.

LEMMA 8. *The set F is nowhere dense.*

The set F is closed and would contain inner points if the lemma were false. But if the fixed points of H contain inner points, H is the identity† and this case has been excluded from consideration.

LEMMA 9. *If O is an open connected set which is invariant under $T(x; t)$, then $O-O\cdot F$ is arcwise connected.*

Suppose that $O-O\cdot F=A+B$ where A and B are disjoined non-vacuous sets such that $A\overline{B}+\overline{A}B=0$. Since $O-O\cdot F$ is open, A and B must be open. The set O is transformed homeomorphically into itself by H, and furthermore $H(A)=A$, $H(B)=B$, and $H(O\cdot F)=O\cdot F$. A new transformation K of O into itself will now be defined. If $x\epsilon B+O\cdot F$, $K(x)=H(x)$; if $x\epsilon A$, $K(x)=x$. The transformation K is a homeomorphism of A into itself. But the set of fixed points of K contains inner points. Since K has period 2 and since O is locally euclidean, it follows from Newman's theorem that every point of O is fixed under K. This contradicts the definition of K and hence $O-O\cdot F$ must be connected. That it is arcwise connected is now immediate, since $O\cdot F$ is closed in O.

It follows at once from this lemma that M is connected.

LEMMA 10. *If q is any fixed point there is a connected open set O including q which has an arbitrarily small diameter and is invariant.*

Take a sphere containing q and let O be the set swept out by this sphere under the action of $T(x; t)$. The conditions on $p(x)$ imply that the diameter of O may be made small by properly choosing the sphere about q.

LEMMA 11. *Every fixed point is arcwise accessible from M.*

This is a consequence of the two preceding lemmas.

PRINCIPAL LEMMA 2. *The set $L(M)$ in E^* is homeomorphic to the euclidean plane.*

Let J denote any simple closed curve of $L(M)$. The set $L^{-1}(J)$ is a torus by Lemma 6, and $\overline{E}-L^{-1}(J)$ is the sum of two connected domains O_1 and O_2, say, such that $\overline{O}_i-O_i=L^{-1}(J)$, $i=1, 2$, by the Alexander Duality Theorem. Since $L^{-1}(J)$ is invariant, it is clear that O_1 and O_2 are invariant. By Lemma 9,

† See Newman, Quarterly Journal of Mathematics, vol. 2 (1931), p. 1.

$O_i - O_i \cdot F$ is connected. It is clear that $L(O_1 - O_1 \cdot F)$ has no point in common with $L(O_2 - O_2 \cdot F)$ and that $L(M) - J$ is equal to the sum of these two sets; that is, J separates $L(M)$. We have, in fact, shown the stronger result that J separates E^*. Now if G is any arc of J, $L^{-1}(G)$ is a closed proper subset of $L^{-1}(J)$. Therefore $\overline{E} - L^{-1}(G)$ is open and connected, and $\overline{E} - L^{-1}(G) - F$ is connected. But $L[\overline{E} - L^{-1}(G) - F] = L(M) - G$. Therefore G does not separate $L(M)$.

To complete the proof it is necessary to show† that one of the domains of $L(M) - J$ is compact in $L(M)$ and that the other is not. One of the domains, we may suppose it is O_2, contains P_∞. It is clear that $L(O_2 - O_2 \cdot F)$ is not compact in $L(M)$. Then we have left to show that O_1 does not contain points of F: in this event $L(O_1)$ *is* compact in $L(M)$. Since J is accessible‡ from each of its domains, if $L(O_1)$ contains a point of $L(F)$, E^* must contain an arc $a_1a_0a_2$ such that $a_1 + a_2$ is in $L(F)$, a_0 is in J, the open arc a_1a_0 is in $L(O_1 - O_1 \cdot F)$, and the open arc a_0a_2 is in $L(O_2 - O_2 \cdot F)$. The sets $L^{-1}(a_1a_0)$ and $L^{-1}(a_2a_0)$ are 2-cells by Lemma 5 and it is clear that their sum is a 2-sphere S_2 which has in common with the torus $L^{-1}(J)$ the simple closed curve $L^{-1}(a_0)$. By the Duality Theorem, S_2 separates E into exactly two open connected domains G_1 and G_2. Now $L^{-1}(J - a_0)$ is connected and must lie entirely in one or the other of these sets, say in G_1. This implies that there cannot be points of G_2 in both O_1 and O_2. Since the points of $L^{-1}(a_1)$ and of $L^{-1}(a_2)$ (points of S_2) must be accessible from G_2, this is a contradiction.

8. We shall now prove the following lemma:

LEMMA 12. *If $\alpha\beta$ is an arc in $L(M)$ and if qr and $q'r'$ are two true sections of $L^{-1}(\alpha\beta)$, there exists a real continuous function $h(x)$ defined everywhere on qr such that $T[x; h(x)]$ is a homeomorphic mapping of qr on $q'r'$.*

Let a denote any point of $\alpha\beta$, x and x' the corresponding points of qr and $q'r'$. It is clear that the correspondence $x' \longleftrightarrow x$ engenders a homeomorphism.

For a definite point x let t be a real number such that $T(x; t) = x'$. Let x_n denote any sequence of points of qr converging to x. Assign to each x_n a number t_n, $|t_n| \leqq \frac{1}{2}p(x_n)$, such that $T(x_n; t + t_n) = x_n'$. At least one such number exists and at most two. We may suppose our sequence of points such that t_n converges to a limit t'. Then $T(x_n; t + t_n) \to T(x; t + t')$ which is equal to $T(x'; t')$. On the other hand $x_n' \to x'$ so that $T(x; t') = x'$, from which we may conclude that $t' = mp(x')$ where m is some integer or zero. Since $p(x)$ is continuous at x', t' cannot exceed $\frac{1}{2}p(x')$ in absolute value, and m must be zero. It follows that for almost all x_n only one choice of t_n is possible and we have

† L. Zippin, *On continuous curves*, American Journal of Mathematics, vol. 52 (1930), pp. 331–350.

‡ L. Zippin, loc. cit.

seen that this choice must be continuous at x. It can now be seen that there must exist a sufficiently small subarc $q''r''$ of qr containing x as an inner point such that a real continuous function $h''(x)$ may be defined on it for which $T[x; h''(x)]$ maps $q''r''$ homeomorphically on the corresponding subarc of $q'r'$. We need only choose an arc so small that the numbers t_n assigned above are actually *less* in absolute value than one-half the corresponding period.

The arc qr may be expressed as a sum of non-overlapping subarcs $q_0q_1, q_1q_2, \cdots, q_{k-1}q_k$, such that there is defined on $q_{i-1}q_i$ a function $h_i(x)$ of the required sort. Now the values of h_1 and h_2 at q_1 can differ only by an integral multiple m of the period at that point. Then $h_2' = h_2 + mp(x)$ is a new function defined on q_1q_2 which is of the required type, and this new function agrees with h_1 at q_1. The function which is equal to h_1 on q_0q_1 and equal to h_2' on q_1q_2 is of the required type over $q_0q_1 + q_1q_2$. We can continue in this manner, and after k steps we arrive at a function defined over qr which satisfies the lemma.

LEMMA 13. *If S is a closed 2-cell in $L(M)$ there exists a true section of $L^{-1}(S)$.*

We can express S as the sum of a finite number of closed 2-cells $A_1, \cdots, A_n$, of small diameter such that $A^i = \sum_1^i A_j$ is itself a closed 2-cell and the intersection of A^k and A_{k+1} is an arc J_{k+1}. Since $L^{-1}(S)$ admits true sections locally by Lemma 3, it may be assumed that to each A_i there corresponds a true section B_i of $L^{-1}(A_i)$.

Let $B^1 = B_1$, and suppose we have constructed a true section B^k of $L^{-1}(A^k)$. Let J be the image of J_{k+1} in B^k and J' its image in B_{k+1}. Let $h(x)$ be the function defined on J' in the preceding lemma, and let $H(x)$ be any continuous extension of this function over all of A_k. Then $B^{k+1} = B^k + T[x; H(x)]$ where x is allowed to vary over B_{k+1}. This set is seen to be a true section of $L^{-1}(A^{k+1})$. The induction is complete and Lemma 13 is proved.

It follows from this lemma that $L^{-1}(S)$ can be mapped homeomorphically into the product of a circular disk D and a circle C in such a way that the path curves of $L^{-1}(S)$ are carried into the sets $d \times C$ where $d \in D$. Hence if S is a closed 2-cell in M and a is a point of S, $L^{-1}(a)$ cannot bound in $L^{-1}(S)$.

9. The fact just proved will be useful in the following lemma:

LEMMA 14. *If a is a point of $L(M)$, $L^{-1}(a)$ cannot bound in M.*†

If $L^{-1}(a)$ bounds in M, it must bound a singular 2-chain K which is contained in M. The transformation L carries K into a closed compact subset of $L(M)$ which we may suppose to be contained in a closed 2-cell S where

† This is very similar to a theorem due to Seifert, Acta Mathematica, vol. 60, p. 224.

$S \subset L(M)$. But now $L^{-1}(a)$ bounds in $L^{-1}(S)$, which from the remark above cannot happen.

Now since any simple closed curve bounds in $\overline{E}$ even after the removal of a finite number of points, it is a corollary to the lemma that F is an *infinite* set.

LEMMA 15. *For every point z in F, $F-z$ is connected.*

If z is a point for which the lemma is false, F may be expressed as the sum of two closed (in F) sets X and Y whose intersection is a subset of z (possibly vacuous). If x and y denote points of X and Y respectively and if $L(x)=a$ and $L(y)=b$, there is an arc ab in E^* which contains no points of $L(F)$ other than a and b. This follows from the accessibility of points of F. Now $L^{-1}(ab)$ is a 2-sphere and bounds in $\overline{E}-z$ by the Alexander Duality Theorem. On the other hand if a_1 is an inner point of ab, $L^{-1}(a_1)$ bounds in the complement of X and in the complement of Y. This situation, however, contradicts corollary W^i of Alexander (loc. cit.), and this contradiction establishes the lemma.

Since F contains infinitely many points and is connected on the removal of any one of them, it must be connected.

PRINCIPAL LEMMA 3. *$L(\overline{E})$ is homeomorphic to a closed half-plane.*

If we let x and y denote any two distinct points of F, and if a, b, and ab are used as in the preceding lemma, the 2-sphere $L^{-1}(ab)$ separates E. Since it is an invariant surface we can conclude that the arc ab separates E^*. A proper subset of this arc gives rise to a proper subset of $L^{-1}(ab)$ and this does not separate E. We conclude that the arc ab separates E^* *irreducibly*.† But it is further clear that each domain of $E-L^{-1}(ab)$ contains at least one orbit, and since this bounds in its domain, we see that each domain must contain fixed points. Therefore the points x and y must separate the set F, and the points a and b must separate $L(F)$. Therefore $L(F)$ is a continuum without cut points which is separated by every pair of its points. It follows that $L(F)$ must be a simple closed curve.‡ Finally we have the fact that E^* is a closed 2-cell with $L(F)$ as its boundary.

It is clear that $E^*-L(P_\infty)$ which is identically $L(E)$ is a closed half-plane, the points $L(F)$ corresponding to the boundary line.

10. We are in a position to bring our entire argument to a close. We shall choose a Cartesian coordinate system (z, r) in $L(E)$, $-\infty < z < \infty$, $0 \leqq r < \infty$, so that $L(F)$ is the z-axis, $r=0$. This is possible by the preceding

† L. Zippin, *Characterization of the closed 2-cell*, American Journal of Mathematics, vol. 55 (1933), pp. 207–217.

‡ R. L. Wilder, *Concerning simple continuous curves*, American Journal of Mathematics, vol. 53 (1931), pp. 39–55; Corollary, p. 48.

lemma. Now $L(M)$ is the point set $r>0$. We can obviously express this set as the sum of a countable number of "rectangles" such that the sum of any finite number of consecutive rectangles is a closed 2-cell. Corresponding to each of these there is a true section in M, by Lemma 13, and these can be "fitted" to each other by the construction in that lemma, where the induction must be carried indefinitely. Therefore M possesses a true section M', and M' is homeomorphic to $L(M)$. The homeomorphism is given by the function L. The function L is certainly a homeomorphism between F and $L(F)$, since each point of F is its own orbit. Then it is clear that L maps $M'+F$ homeomorphically into $L(E)$. For if a sequence of points of $L(M)$ converges to a point of $L(F)$, the corresponding orbits must converge to the fixed point in E, and therefore the corresponding points of M', which lie on these orbits, must converge to it.

We may now transfer to $E''=M'+F$ the coordinate system (z, r). We shall extend this to a cylindrical coordinate system in all of E. Consider now an arbitrary point x of E, our original three-space. It belongs to an orbit O_x which intersects E'' in a single point, say x'', with coordinates (z'', r''), say. If $r''=0$, x'' is a fixed point and $x=x''$. Then the "θ-coordinate" of x shall be indeterminate. If $r''>0$, there is a unique t'', $0\leqq t''<p(x'')=p(x)$, such that $T(x'', t'')=x$. The θ coordinate of x shall be $2\pi t''/p(x'')$. Conversely, it is clear that any three numbers z, θ, r,

$$-\infty < z < \infty, \qquad 0 \leqq \theta \leqq 2\pi,$$

determine a unique point of E. With this coordinate system established in E we may take E to be the "reference-space" R of §1, and the homeomorphism of that section to be identity. The function $F(z, r)$ is now $1/p(z, r)$, $p(z, r)$ being the value of the period function at the point (z, r) of E''. We have seen that this function is independent of θ and is continuous at the moving points, i.e., points for which $r>0$.

PRINCETON UNIVERSITY, AND
INSTITUTE FOR ADVANCED STUDY,
PRINCETON, N. J.

THE THEORY OF REPRESENTATIONS FOR BOOLEAN ALGEBRAS*

BY

M. H. STONE

INTRODUCTION

Boolean algebras are those mathematical systems first developed by George Boole in the treatment of logic by symbolic methods† and since extensively investigated by other students of logic, including Schröder, Whitehead, Sheffer, Bernstein, and Huntington.‡ Since they embody in abstract form the principal algebraic rules governing the manipulation of classes or aggregates, these systems are of technical interest to the mathematician quite as much as to the logician. It is thus natural to suppose that a study of Boolean algebras by the methods of modern algebra will prove fruitful of important and useful results. Indeed, if one reflects upon various algebraic phenomena occurring in group theory, in ideal theory, and even in analysis, one is easily convinced that a systematic investigation of Boolean algebras, together with still more general systems, is probably essential to further progress in these theories.§ The writer's interest in the subject, for example, arose in connection with the spectral theory of symmetric transformations in Hilbert space and certain related properties of abstract integrals. In the actual development of the proposed theory of Boolean algebras, there emerged some extremely close connections with general topology which led at once to results of sufficient importance to confirm our a priori views of the probable value of such a theory.‖

In the present paper, which is one of a projected series, we shall be concerned primarily with the problem of determining the representation of a

* Presented to the Society (in part), February 25, 1933; see abstract 39-3-86. Received by the editors October 10, 1935.

† See *The Mathematical Analysis of Logic*, 1847, and *An Investigation of the Laws of Thought*, 1854.

‡ For bibliographical information, see Huntington, these Transactions, vol. 35 (1933), pp. 274–304, especially pp. 274–275.

§ Dedekind's observations and technical contributions, in Mathematische Annalen, vol. 53 (1900), pp. 371–403, relative to the occurrence of such general systems did not immediately provoke general interest. Recently, a considerable amount of work along the general lines laid down by Dedekind has appeared: see, for instance, Fritz Klein, Mathematische Annalen, vol. 105 (1931), pp. 308–323; Garrett Birkhoff, Proceedings of the Cambridge Philosophical Society, vol. 29 (1933), pp. 441–464; O. Ore, Annals of Mathematics, (2), vol. 36 (1935), pp. 406–437.

‖ See Stone, Proceedings of the National Academy of Sciences, vol. 20 (1934), pp. 197–202.

given Boolean algebra by algebras of classes, aggregates, or combinations. It is natural to surmise that the problem always has a solution leading to the construction of an algebra of classes isomorphic to the given Boolean algebra. Such a result is a precise analogue of the theorem that every abstract group is represented by an isomorphic group of permutations. Here we shall establish the validity of this surmise and, in addition, shall characterize all possible algebras of classes homomorphic to a given Boolean algebra. It is a curious fact that these results are considerably more recondite than the corresponding theorems for abstract groups: the elements of the representative classes must be taken as certain classes of elements in the given Boolean algebra (in particular, as the prime ideals in the algebra), whereas the elements of the permutations representing an abstract group are taken as elements of the group itself; and the existence of prime ideals, in terms of which the representation is constructed, can apparently be established in general only by an appeal to the Zermelo hypothesis.

The observation that Boolean algebras can be regarded as special instances of the systems known as abstract rings enables us to apply the concepts and results of modern algebraic theory directly to the purposes of the present paper.* Here we shall show in detail that Boolean algebras are identical with those rings with unit in which every element is idempotent.† In this identification, ring addition and multiplication correspond abstractly to the formation of the union (modulo 2) and intersection, respectively, of classes. The union (modulo 2), or symmetric difference, of two classes is the class of objects belonging to one or the other, but not to both, of those classes; and is thus familiar to combinatorial topologists. On algebraic grounds it is convenient to admit rings other than those which possess units. We shall therefore take as the central theme of this paper not merely Boolean algebras, but, more generally, rings in which every element is idempotent, designating the latter systems as Boolean rings or generalized Boolean algebras.

The paper falls naturally into four parts or chapters. The first deals with the formal algebraic properties of Boolean rings; the second with subrings, ideals, and homomorphisms; the third with the structure of Boolean rings with elements which are given as abstract classes; and the fourth with the representation theory. In general we shall limit our investigations to those topics which are essential for an adequate understanding of the algebraic

* For general information concerning modern algebraic developments, we refer to B. L. van der Waerden, *Moderne Algebra*, vol. 1, Berlin, 1930, Chapter 3 of which deals particularly with the definition and basic properties of abstract rings. We shall assume that the reader has a general knowledge of this material.

† See Stone, Proceedings of the National Academy of Sciences, vol. 21 (1935), pp. 103–105; Gegalkin, Matematicheskii Sbornik, vol. 35 (1928), pp. 311–373.

aspects of the representation theory, leaving until another occasion a deeper study of the classification of ideals and the introduction of the concepts of general topology. A more complete survey of the contents of the present paper is the following: Chapter I, Formal algebraic properties of Boolean rings: §1, Direct discussion of Boolean rings; §2, Connections with Boolean algebras; §3, Special Boolean rules; §4, Special elements. Chapter II, Subrings, ideals, and homomorphisms: §1, Subrings and their combinations; §2, Ideals and their combinations; §3, A classification of ideals; §4, Prime ideals; §5, Congruences, ideals and homomorphisms; §6, Direct sums. Chapter III, Algebras of classes: §1, The construction of algebras of classes; §2, Reduction and equivalence; §3, The analysis of algebras of classes; §4, An illustration. Chapter IV, Representation theory: §1, General remarks; §2, Existence and divisibility properties of prime ideals; §3, The perfect representation.

Chapter I. Formal algebraic properties of Boolean rings

1. **Direct discussion of Boolean rings.** In this section we shall consider the elementary facts relating to rings in which every element is idempotent, leaving for the remaining sections of the chapter the study of their connections with Boolean algebras and of certain special relations and elements. By a ring we mean a system with double composition, the operations being called addition and multiplication and denoted here by the usual symbols $+$ and $\cdot$ (the latter commonly being suppressed in writing down products), subject to the following laws: addition is commutative and associative, multiplication is associative and both left- and right-distributive with respect to addition, and the equation $x+a=b$ has a solution for arbitrary a and b. We do not assume that multiplication is commutative or that a ring contains more than one element. It is well known that in any ring the solution 0 of the equation $x+a=a$ is independent of a and satisfies the relations $a+0=0+a=0$, $0a=a0=0$; that the solution $-a$ of the equation $x+a=0$ is unique; and that the solution of the equation $x+a=b$ is unique and is given by $x=b+(-a)$, commonly written $b-a$. We now lay down the following formal definition:

Definition 1. *A ring in which every element is idempotent, satisfying the law $aa=a$, is called a Boolean ring.*

Our first result is embodied in the following theorem.

Theorem 1. *A Boolean ring is necessarily commutative; obeys the two equivalent laws $a+a=0$, $a=-a$; and necessarily contains divisors of 0 if it contains more than two elements. Every Boolean ring A can be imbedded in a Boolean*

ring B which possesses a unit element, in such a manner that B is unique in the following sense: if C is a Boolean ring with unit containing A, then C contains also a Boolean ring B isomorphic to B and containing A. A finite Boolean ring necessarily possesses a unit and has a cardinal number which is a power of 2.*

In this, as in all subsequent discussions, we may use the familiar rules governing the ring operations without going into complete detail. Using such rules, we see that in a Boolean ring

$$a + b = (a + b)(a + b) = (a + b) + (ba + ab)$$

and hence that $ba+ab=0$. If we put $b=a$ in the latter relation, we find at once that $a+a=0$, or, equivalently, $a=-a$. Using this result, we conclude that $ba=-(ab)=ab$, thus establishing the commutative law for multiplication. The special rules which we have now demonstrated will henceforth be used in our discussions without explicit reference. In a Boolean ring with more than two elements, we can choose a and b so that $a\neq 0$, $b\neq 0$, $a\neq b$. If $ab=0$, then a and b are both divisors of 0. On the other hand, if $ab\neq 0$, then ab and $a+b$ are both divisors of 0: for $a+b=0$ would imply $a=-b=b$, contrary to hypothesis; and $ab(a+b)=aab+abb=ab+ab=0$. A Boolean ring with one or with two elements obviously cannot contain divisors of 0, every product in such a ring either containing 0 as a factor or reducing, by the law of idempotence, to an element other than 0.

In discussing the possibility of imbedding a Boolean ring A in a Boolean ring B with unit, we may disregard the trivial case where A has a unit and B coincides with A. When A has no unit, we construct B by the adjunction of suitable elements. The construction can be carried out even when A has a unit and always produces B as a proper superclass of A. We first provide an abstract element ϵ, distinct from those of A, and define

$$\epsilon\epsilon = \epsilon, \quad \epsilon a = a\epsilon = a, \quad \epsilon + 0 = 0 + \epsilon = \epsilon, \quad \epsilon + \epsilon = 0,$$

observing that the elements 0 and ϵ constitute a two-element Boolean ring. We then consider the ordered pairs (a, α) where a is in A and $\alpha=0$ or $\alpha=\epsilon$, defining the operations of addition and multiplication upon them by the rules

$$(a, \alpha) + (b, \beta) = (a + b, \alpha + \beta),$$
$$(a, \alpha)(b, \beta) = (ab + \alpha b + a\beta, \alpha\beta).$$

It is easily verified that under these operations the class of pairs (a, α) is a Boolean ring with $(0, \epsilon)$ as its unit. Instead of giving all the calculations in detail, we shall discuss only one or two steps. Passing over the commutative

and associative laws for addition and multiplication, we turn to the demonstration of the left-distributive law for multiplication: we have

$$\begin{aligned}(a, \alpha)[(b, \beta) + (c, \gamma)] &= (a, \alpha)(b + c, \beta + \gamma)\\ &= (a(b + c) + \alpha(b + c) + a(\beta + \gamma), \alpha(\beta + \gamma))\\ &= ((ab + \alpha b + a\beta) + (ac + \alpha c + a\gamma), \alpha\beta + \alpha\gamma)\\ &= (ab + \alpha b + a\beta, \alpha\beta) + (ac + \alpha c + a\gamma, \alpha\gamma)\\ &= (a, \alpha)(b, \beta) + (a, \alpha)(c, \gamma)\end{aligned}$$

by appropriate combination of the properties of A with the easily checked relations $\alpha(b+c)=\alpha b+\alpha c$, $a(\beta+\gamma)=a\beta+a\gamma$, $\alpha(\beta+\gamma)=\alpha\beta+\alpha\gamma$. The right-distributive law then follows from this by the use of the commutative law for multiplication. We now see that the equation $(x, \xi)+(a, \alpha)=(b, \beta)$ has the solution $(x, \xi)=(b+a, \beta+\alpha)$ since

$$(b + a, \beta + \alpha) + (a, \alpha) = (b + a + a, \beta + \alpha + \alpha) = (b, \beta).$$

Thus the pairs (a, α) constitute a commutative ring. The law of idempotence is established as follows:

$$\begin{aligned}(a, \alpha)(a, \alpha) &= (aa + \alpha a + a\alpha, \alpha\alpha) = (a + (\alpha + \alpha)a, \alpha)\\ &= (a + 0a, \alpha) = (a, \alpha).\end{aligned}$$

To show that the Boolean ring thus obtained has $(0, \epsilon)$ as its unit, we note the relations

$$(0, \epsilon)(a, \alpha) = (0a + \epsilon a + 0\alpha, \epsilon\alpha) = (a, \alpha).$$

As in any commutative ring, the unit is unique and is a right- as well as left-unit. The pairs $(a, 0)$ obviously constitute a Boolean ring isomorphic to A in accordance with the equations

$$\begin{aligned}(a, 0) + (b, 0) &= (a + b, 0),\\ (a, 0)(b, 0) &= (ab + 0b + a0, 00) = (ab, 0).\end{aligned}$$

To construct B, we now replace each pair $(a, 0)$ by the corresponding element a without disturbing in any other way the operations defined over the class of pairs (a, α). Evidently B is a Boolean ring with unit isomorphic to the ring of pairs (a, α); and it contains A. If C is a Boolean ring with unit e containing A, we set up an isomorphism between B and a subring B^* of C as follows. The elements of B correspond in a one-to-one manner with the pairs (a, α), by construction. We can easily set up a one-to-one correspondence between the pairs (a, α) and certain elements of C by requiring that $(a, \alpha)\longleftrightarrow a+f(\alpha)$, where a is in A and $f(0)=0$, $f(\epsilon)=e$. The fact that e be-

longs to C but not to A renders the indicated correspondence biunivocal, since $a+f(\alpha)=b+f(\beta)$ implies $f(\alpha)+f(\beta)=a+b\neq e$, $f(\alpha)+f(\beta)=0$, $f(\alpha)=f(\beta)$, $a=b$. In view of the relations $f(\alpha+\beta)=f(\alpha)+f(\beta)$, $f(\alpha\beta)=f(\alpha)f(\beta)$, which are readily checked, the correspondence between the pairs (a, α) and the elements $a+f(\alpha)$ is an isomorphism: for $(a, \alpha)\longleftrightarrow a+f(\alpha)$, $(b, \beta)\longleftrightarrow b+f(\beta)$ imply

$$\begin{aligned}(a, \alpha) + (b, \beta) &= (a + b, \alpha + \beta) \longleftrightarrow (a + b) + f(\alpha + \beta)\\ &= (a + f(\alpha)) + (b + f(\beta)),\\ (a, \alpha)(b, \beta) &= (ab + \alpha b + a\beta, \alpha\beta) \longleftrightarrow ab + \alpha b + a\beta + f(\alpha\beta)\\ &= ab + f(\alpha)b + af(\beta) + f(\alpha)f(\beta)\\ &= (a + f(\alpha))(b + f(\beta)).\end{aligned}$$

The elements $a+f(\alpha)$ thus constitute a Boolean ring B^* isomorphic to B, containing the elements of A (since $a=a+f(0)$), and contained in C.

If A is a finite Boolean ring, we determine its unit explicitly as a symmetric function of its elements. Denoting by $p(a_1, \cdots, a_n)$ the sum of the n elementary symmetric functions of the elements $a_1, \cdots, a_n$, we observe the identity

$$\begin{aligned}p(a_1, \cdots, a_n) = a_k &+ p(a_1, \cdots, a_{k-1}, a_{k+1}, \cdots, a_n)\\ &+ a_k p(a_1, \cdots, a_{k-1}, a_{k+1}, \cdots, a_n),\end{aligned}$$

valid for any commutative ring. In a Boolean ring A with exactly N elements, we show that $p(a_1, \cdots, a_N)$ is a unit: for, if a_k is one of the elements, the above identity leads, in combination with the peculiar properties of A, to the relation

$$p(a_1, \cdots, a_N)a_k = a_k + 2p(a_1, \cdots, a_{k-1}, a_{k+1}, \cdots, a_N)a_k = a_k.$$

If A is a finite Boolean ring with exactly N elements, we show further that N is a power of 2. Indeed, we shall establish the more general result that an additive abelian group of N elements in which $a+a=0$ has the property that N is a power of 2. For $N=1$, the desired result is obvious. Let us suppose that this result has been extended to those cases where $N\leqq 2^M$. If now $N\leqq 2^{M+1}$ we may suppose also that known cases are excluded by subjecting N to the inequality $N>2^M$. The given group G contains an element a other than 0, the two elements a and 0 constituting a normal subgroup $\mathfrak{g}$ by virtue of the relations $0+0=a+a=0$, $a+0=0+a=a$. The quotient group $G/\mathfrak{g}$ is homomorphic to G and therefore obeys the law $a+a=0$. Since its order is $N/2$ and therefore at most equal to 2^M, we conclude that $N/2=2^m$, $N=2^{m+1}$. The inequalities satisfied by N show that $N=2^{M+1}$. By induction the desired

result holds for all finite groups of the indicated type and hence, in particular, for finite Boolean rings.

2. **Connections with Boolean algebras.** We shall prove in the present section that the fundamental properties of Boolean rings, as stated in Definition 1, in reality constitute a new set of postulates for Boolean algebras and certain simple generalizations thereof, these systems having hitherto been characterized by the postulation of different fundamental properties governing operations among which the ring-addition of the preceding section did not appear. Our task thus consists in setting up suitable relations between the respective operations chosen as fundamental in various sets of postulates, including the set suggested by Definition 1, and in establishing the interdeducibility of the postulates of these different sets. Another task which we might undertake is that of eliminating possible redundancies from the set of ring-postulates, as here extended to include the law of idempotence. We shall not make any serious attempt to carry through the necessary investigations in the present paper, such an enterprise being of secondary importance for our immediate purposes. What is of primary importance here is the identification of the abstract algebras arising from logic and the theory of classes with systems amenable to the methods developed by modern algebraists, namely, with those special rings which we have termed Boolean rings in anticipation of the detailed results of the present section.

We shall first consider a recent set of postulates for Boolean algebras due to Huntington.† The operations in terms of which the postulates are framed are a binary operation, which we shall denote by $\vee$ and which corresponds to logical addition and to the formation of the union for classes, and a unary operation, which is denoted by the prime $'$ and which corresponds to logical negation and to the formation of the complement for classes. These operations are assumed to apply without restriction to elements of the system, yielding elements of the system. Huntington's postulate requiring that the system contain at least two elements plays no part in his deductions and will be suppressed here so as to admit one-element Boolean algebras as well as one-element rings. In stating the connection between Huntington's characterization of Boolean algebras and the properties of Boolean rings, we shall designate several propositions by the numbers attached to them in Huntington's paper, thus facilitating comparison.

THEOREM 2. *If A is a Boolean ring with unit e, the introduction of a binary operation $\vee$ and a unary operation $'$ through the equations*

† Huntington, these Transactions, vol. 35 (1933), pp. 274–304, 557–558. The set in question is discussed on pp. 280–286, 557–558, a serious redundancy in the original set being eliminated on the two pages last cited.

(1) $a \vee b = a + b + ab$, (2) $a' = a + e$

converts A into an algebraic system B in which

(4.3) $a \vee b = b \vee a$, (4.4) $a \vee (b \vee c) = (a \vee b) \vee c$,

(4.6) $(a' \vee b')' \vee (a' \vee b)' = a$,

the old operations being expressed in terms of the new through the equations

(6) $a + b = ab' \vee a'b = (a' \vee b'')' \vee (a'' \vee b')'$,

(7) $ab = (a' \vee b')'$.

On the other hand, if B is an algebraic system obeying the laws (4.3), (4.4) *and* (4.6), *then B is a Boolean algebra; and the introduction of new operations through the equations* (6) *and* (7) *converts B into a Boolean ring A with unit* $e = a \vee a'$ *and zero* $0 = e' = (a \vee a')'$, *the old operations being expressed in terms of the new through the equations* (1) *and* (2) *above.*†

This theorem clearly serves to identify Boolean rings with unit and Boolean algebras, as characterized by Huntington's postulates. In view of Theorem 1, a Boolean ring without unit can be regarded as imbedded in one which has a unit. Hence the present theorem reveals the essential nature of all Boolean rings. In particular, it shows that the operation of addition in a Boolean ring corresponds abstractly to the operation of forming the symmetric difference or union (modulo 2) of classes, as indicated by the relation (6); and it shows similarly that the operation of multiplication corresponds to the operation of forming the intersection of classes, as indicated in the relation (7).

If A is any Boolean ring, either with or without unit, we may write $a \vee b = a + b + ab = p(a, b)$, in terms of the symmetric function introduced in the proof of Theorem 1. We then have

$$a \vee (b \vee c) = p(a, p(b, c)) = a + p(b, c) + ap(b, c) = p(a, b, c)$$
$$= p(p(a, b), c) = (a \vee b) \vee c.$$

Thus the operation introduced in (1) has properties (4.3) and (4.4), expressing the commutative and associative laws. If A has a unit e, we observe that the operation introduced in (2) has the properties

$$a'' = (a + e) + e = a + (e + e) = a + 0 = a,$$
$$(a \vee b)' = (a + b + ab) + e = (a + e)(b + e) = a'b'.$$

With their aid we can establish (4.6), (6), and (7), as follows:

† Stone, Proceedings of the National Academy of Sciences, vol. 21 (1935), pp. 103–105.

$$
\begin{aligned}
(a' \vee b')' \vee (a' \vee b)' &= a''b'' \vee a''b' = ab \vee ab' \\
&= ab + ab' + (ab)(ab') \\
&= ab + a(b + e) + ab(b + e) \\
&= ab + ab + a + ab + ab = a, \\
(a' \vee b'')' \vee (a'' \vee b')' &= (a' \vee b)' \vee (a \vee b')' = a''b' \vee a'b'' = ab' \vee a'b \\
&= a(b + e) + (a + e)b + ab(a + e)(b + e) \\
&= ab + a + ab + b + (a + a)(b + b) = a + b, \\
(a' \vee b')' &= a''b'' = ab.
\end{aligned}
$$

Thus the introduction of the new operations $\vee$ and $'$ converts A into a Boolean algebra B in the sense of Huntington's postulates.

On the other hand, if B is a Boolean algebra in that sense, Huntington has shown that the following propositions are valid in B:

(4.10) $a'' = a$, (4.11) $a \vee a' = b \vee b'$, (4.5) $a \vee a = a$,

(4.16) $ae = a$, (4.22) $a \vee e = e$,

(4.15) $0 \vee a = a$, (4.23) $a0 = 0$,

(4.18) $ab = ba$, (4.19) $(ab)c = a(bc)$, (4.34) $a(b \vee c) = ab \vee ac$,

where e is the element $a \vee a'$, unique in accordance with (4.11), 0 is the element e', and $ab = (a' \vee b')'$ in accordance with (7). On noting that $aa' = (a' \vee a'')' = e' = 0$, we see that

$$
\begin{aligned}
a + b &= ab' \vee a'b = ba' \vee b'a = b + a, \\
a + (b + c) &= a(bc' \vee b'c)' \vee a'(bc' \vee b'c) = a(bc')'(b'c)' \vee a'bc' \vee a'b'c \\
&= a(b' \vee c)(b \vee c') \vee a'bc' \vee a'b'c = ab'c' \vee abc \vee a'bc' \vee a'b'c \\
&= ca'b' \vee cab \vee c'ab' \vee c'a'b = c + (a + b) \\
&= (a + b) + c
\end{aligned}
$$

by virtue of the propositions stated above. We see further that

$$
\begin{aligned}
ab + ac &= ab(ac)' \vee (ab)'(ac) = ab(a' \vee c') \vee (a' \vee b')(ac) \\
&= abc' \vee ab'c = a(bc' \vee b'c) = a(b + c).
\end{aligned}
$$

Thus the operations $+$ and $\cdot$ are both commutative and associative and the second is left- and right-distributive with respect to the first. If we can show that the equation $x + a = b$ has a solution, we can therefore assert that the introduction of these operations converts B into a commutative ring A. Now we evidently have $a + a = aa' \vee a'a = 0$ and $a + 0 = a0' \vee a'0 = a0' = ae = a$. Consequently, we find that $x = b + a$ is a solution in accordance with the rela-

tions $(b+a)+a=b+(a+a)=b+0=b$. Since the indicated solution reduces to $x=0$ when $b=a$, we have also identified 0 as the zero element of the ring A into which B is converted. The proposition (4.16) identifies e as the unit of A. In order to show that A is a Boolean ring, we have only to observe the equations $aa=(a' \vee a')'=a''=a$. To complete the proof of the theorem we must show finally that relations (1) and (2) are valid. Since we have

$$\begin{aligned} a + e &= ae' \vee a'e = a0 \vee a' = a', \\ a \vee b &= (a \vee b)'' = (a'b')' = (a + e)(b + e) + e \\ &= ab + a + b + e + e \\ &= a + b + ab, \end{aligned}$$

the desired results are established.

The postulates of Huntington which have just been discussed are not the only ones in terms of which Boolean algebras may be characterized. Indeed, the sets of postulates which involve only the operations corresponding to the formation of the union and intersection of classes are quite numerous, should one wish to confine himself to those postulates which are perhaps the most natural as well as the most familiar. It may therefore be of some interest if we establish a theorem relating Boolean rings to Boolean algebras as characterized by at least one such set of postulates. In any event, the possibility of extending such a relation so as to obtain a characterization of all Boolean rings, both with and without unit, in terms of union and intersection surely deserves consideration. We shall therefore proceed to discuss two sets of postulates which we have given elsewhere, one set characterizing Boolean algebras, the other slightly more general systems which we have called generalized Boolean algebras and shall now identify with Boolean rings.†

THEOREM 3. *If A is a Boolean ring with unit e, then the replacement of the operation $+$ by a new operation $\vee$ defined by the relation*

$$(1) \qquad a \vee b = a + b + ab$$

converts A into a system B with the properties

$$(1_1) \quad a \vee b = b \vee a;$$

$$(3_1) \quad a(b \vee c) = ab \vee ac; \qquad (3_2) \quad (a \vee b)c = ac \vee bc;$$

(4_1) there exists an element 0 such that $a \vee 0 = a$ for every a;

(5) if there exists an element 0 with the property (4_1), then there exists at least one such element 0 to which corresponds a fixed element e such that the equations $x \vee a = e$, $xa = 0$ have a solution for every element a;

† Stone, American Journal of Mathematics, vol. 57 (1935), pp. 703–732.

(6_1) $a \vee a = a$; (6_2) $aa = a$;

where the old operation $+$ *is defined in terms of the new by the relation*

(7) $a+b$ *is a solution, necessarily unique, of the simultaneous equations* $x \vee ab = a \vee b$, $x(ab) = 0$.

Conversely, if B *is a system with the indicated properties* (1_1)–(6_2), *the replacement of the operation* $\vee$ *by the new operation* $+$ *defined by the relation* (7) *converts* B *into a Boolean ring* A *with the elements* 0 *and* e *of* (4_1) *and* (5) *as its zero and unit elements respectively, the old operation* $\vee$ *being expressed in terms of the new by the relation* (1).

If A is a Boolean ring with unit e, we verify properties (1_1)–(6_2) as follows: (1_1) is proved as in Theorem 2; (3_1) is proved by the relations $a(b \vee c) = a(b+c+bc) = ab+ac+abc = ab+ac+(ab)(ac) = ab \vee ac$; (3_2) follows from (3_1) by the commutative law for multiplication, already proved in Theorem 1; the element 0 of A has the property (4_1), since $a \vee 0 = a+0+a0 = a$; for the solution x of the equations given in (5) we may take $x = a+e$, where e is the unit in A, since $(a+e) \vee a = (a+e)+a+(a+e)a = a+e+a+a+a = e$, $(a+e)a = a+a = 0$; (6_1) is proved by the relations $a \vee a = a+a+aa = a+a+a = a$; and (6_2) is the characteristic property of Boolean rings. Thus A is converted into a system B in the indicated manner. The relation (7) is verified by virtue of the equations

$$(a + b)ab = ab + ab = 0,$$

$$(a + b) \vee ab = a + b + ab + (a + b)ab = a + b + ab = a \vee b.$$

The converse part of the theorem is proved, though not explicitly stated, in the paper cited above. The properties (1_1)–(6_2) are there shown to be characteristic of Boolean algebras and the solution of the equations (7) is discussed at length, being denoted by $a \Delta b$ instead of $a+b$. The commutative and associative laws for the operation Δ or $+$, and for multiplication, the distributive laws for multiplication, and the existence of a solution of the equation $x+a = b$, are established in Theorems 26, 34, 13, 15, 38, and 33 respectively. Thus the replacement of the operation $\vee$ by the operation $+$ converts B into a commutative ring A. Property (6_2) identifies A as a Boolean ring; and the existence of a unit is established in Theorem 2 of the cited paper, where it is shown that the element e of (5) has the property $ea = a$. The relation $a+0 = a$, proved in Theorem 29, identifies the element 0 of (4_1) as the zero of the ring A. Finally we establish the relation (1) as follows: by (7) we see that $(a+b)+ab$ must satisfy the equation $x \vee (a+b)ab = (a+b) \vee ab$ and hence must be equal to $(a+b) \vee ab$ since $(a+b)ab = ab+ab = 0$; but, by (7) again, $(a+b) \vee ab = a \vee b$.

THEOREM 4. *If A is a Boolean ring, either with or without unit, the replacement of the operation $+$ by the operation $\vee$ defined by the relation*

$$(1) \qquad a \vee b = a + b + ab$$

converts A into a system B with the properties

$$(1_1) \qquad a \vee b = b \vee a;$$

$$(2_2) \qquad a(bc) = (ab)c;$$

$$(3_1) \qquad a(b \vee c) = ab \vee ac;$$

(4_1) *there exists an element* 0 *such that* $a \vee 0 = a$ *for every* a;

(5_1) *if* $ba = a$, *there exists an element* 0 *with property* (4_1), *independent of* a *and* b, *such that the equations* $x \vee a = b$, $xa = 0$ *have a solution*;

(5_2) *if* $ab = a$, *there exists an element* 0 *with the property* (4_1), *independent of* a *and* b, *such that the equations* $x \vee a = b$, $ax = 0$ *have a solution*;

$$(6_1) \quad a \vee a = a; \qquad (6_2) \quad aa = a;$$

where the old operation $+$ is defined in terms of the new by the relation

(7) $a+b$ *is a solution, necessarily unique, of the simultaneous equations* $x \vee ab = a \vee b$, $x(ab) = 0$.

Conversely, if B is a system with the indicated properties (1_1)–(6_2), *the replacement of the operation $\vee$ by the new operation $+$ defined by the relation* (7) *converts B into a Boolean ring A with the element* 0 *of* (4_1) *as its zero element, the old operation $\vee$ being expressed in terms of the new by the relation* (1).

This theorem serves to identify Boolean rings with those systems which we have termed generalized Boolean algebras.

If A is a Boolean ring, we see from the discussion given under Theorem 3 that the operation $\vee$ defined by (1) has properties (1_1), (3_1), (4_1), (6_1), and (6_2). Property (2_2), the associative law for multiplication, holds in A by the definition of a ring. Thus we have only to examine properties (5_1) and (5_2). In view of the commutative law for multiplication in A, these properties are equivalent, and we may confine our attention to the first. We show that $x = a+b$ satisfies the simultaneous equations of (5_1): $x \vee a = (a+b)+a + (a+b)a = a+b+a+a+a = b$, when $ba = a$; and $xa = (a+b)a = a+a = 0$, when $ba = a$. Hence the replacement of the operation $+$ by the operation $\vee$ converts A into a system B of the indicated type. The proof that the relation (7) is valid is the same as that given under Theorem 3.

The converse part of the theorem is proved, but not explicitly stated, in the paper cited above. In fact, Theorem 55 of that paper shows that the argument used in Theorem 3 above applies equally well to the present system B, save for the part relating to the existence of a unit.

3. **Special Boolean rules.** The results of §§1 and 2 show that the various operations $+$, $\vee$, $\cdot$, and $'$ obey all the formal rules peculiar to appropriate corresponding operations upon classes. Many of these special rules, which we may properly designate as Boolean rules, are stated explicitly in the text. Such, for example, are the commutative and associative laws for each of the operations $+$, $\vee$, and $\cdot$, and the distributive laws for $\cdot$ with respect to $+$ and $\vee$. Indeed, the only important rule which is neither stated nor trivially implied by rules given explicitly is the distributive law for $\vee$ with respect to $\cdot$, a rule which we can prove at once in the following manner:

$$\begin{aligned}(a \vee b)(a \vee c) &= (a + b + ab)(a + c + ac) \\ &= a + ac + ac + ba + bc + bac + ab + abc + abc \\ &= a + bc + abc = a \vee bc.\end{aligned}$$

In the sequel, we shall leave the justification for our use of such special Boolean rules to the reader's recollection of the familiar corresponding rules for operations upon classes or to his personal verification of their validity on the basis of Theorem 1 and the relations $a' = a + e$, $a \vee b = a + b + ab$. It is perhaps desirable that we should point out one essential fact at this place: we raise no presumption that every possible law verifiable for the appropriate operations upon classes is also verifiable in the abstract for the associated operations in Boolean rings. Indeed, the proof that such is the case is one of the central features of the representation theory which we propose to build upon the basis of rules already cited or, at least, readily verifiable therefrom.

It is convenient for us to introduce at this point the abstract relation which corresponds to the relation of class-inclusion and to outline its chief properties. More exhaustive investigations of this topic are to be found elsewhere.† We begin with the formal definition of the indicated relation.

DEFINITION 2. *In a Boolean ring A, the element a is said to be less than or to be contained in the element b, in symbols $a < b$, and the element b is said to be greater than or to contain the element a, in symbols $b > a$, whenever any of the equivalent relations*

$$ab = a, \qquad a \vee b = b, \qquad ab' = 0, \qquad a' \vee b = e$$

is satisfied, the last two being significant if and only if A has a unit e.

The equivalence of the indicated relations is evident if we rewrite them in terms of ring addition and multiplication as

† See Huntington, these Transactions, vol. 5 (1904), pp. 288–309; or Stone, American Journal of Mathematics, vol. 57 (1935), pp. 703–732.

$ab = a, \quad a + b + ab = b, \quad ab + a = 0, \quad (a + e) + b + (ab + b) = e,$

respectively.

The chief properties of this relation are given in the following theorem:

THEOREM 5. *The relation $<$ of Definition 2 obeys the rules*

(1) *$a<b$ and $b<c$ imply $a<c$;*

(2) *$0<a$ for every a, and $a<e$ for every a when the Boolean ring A has a unit e;*

(3) *$a<c$ and $b<d$ imply $ab<cd$, $a \vee b<c \vee d$;*

(4) *$bc=0$ implies $ac=0$ if and only if $a<b$.*

The proof of (1) is obtained as follows: $ab=a$ and $bc=b$ together imply $ac=(ab)c=a(bc)=ab=a$. Property (2) results from the equations $0a=0$, $ae=a$. The proof of (3) follows from the relations, valid when $ac=a$ and $bd=d$,

$$\begin{aligned}(ab)(cd) &= (ac)(bd) = ab,\\ (a \vee b)(c \vee d) &= (a + b + ab)(c + d + cd)\\ &= ac + ad + acd + bc + bd + bcd + abc + abd + abcd\\ &= a + ad + ad + bc + b + bc + ab + ab + ab\\ &= a + b + ab = a \vee b.\end{aligned}$$

We verify (4) in two steps: first, if $a<b$, we see that $ab=a$ and $bc=0$ imply $ac=(ab)c=a(bc)=a0=0$; and, secondly, we see that, if $ac=0$ whenever $bc=0$, the fact that $b(a+ab)=ab+ab=0$ enables us to conclude that $a+ab=a(a+ab)=0$, $ab=a$, and $a<b$.

4. **Special elements.** We shall now turn to the study of certain special elements, other than the zero and unit elements, in Boolean rings, namely, those elements which may be briefly described as minimal non-zero elements. It should be observed that we do not assert the existence of such elements in general; we merely consider what occurs when they do exist. We first lay down some formal definitions.

DEFINITION 3. *A non-zero element a in a Boolean ring A is said to be an atomic element if it has either of the following equivalent properties:*

(1) *$a>b$ implies $b=a$ or $b=0$;*

(2) *$ab=0$ or $ab=a$ for every b.*

The equivalence of properties (1) and (2) is evident from the fact that $ab=b$ when $a>b$ and that $a>ab$ for every b.

DEFINITION 4. *A class $\mathfrak{s}$ of atomic elements is said to be an atomic basis if every non-zero element is the sum of elements in $\mathfrak{s}$.*

DEFINITION 5. *A class $\mathfrak{s}$ of atomic elements is said to be a complete atomic system if $b=0$ is the only element such that $ba=0$ for every a in $\mathfrak{s}$.*

We shall establish several simple theorems concerning systems of atomic elements.

THEOREM 6. *If a and b are atomic elements, then $a=b$ or $ab=0$.*

For $ab\neq 0$ implies both $ab=a$ and $ab=b$ in accordance with Definition 3.

THEOREM 7. *A complete atomic system in a Boolean ring A contains every atomic element in A.*

For an atomic element a, being different from 0, cannot satisfy the relation $ab=0$ for every b in the given system and must therefore be equal to some element b in that system, by virtue of Theorem 6.

THEOREM 8. *If A is a Boolean ring, $\mathfrak{s}$ a complete atomic system in A, and $\mathfrak{s}(b)$ the class of all atomic elements a in $\mathfrak{s}$ such that $ab\neq 0$, then A is isomorphic to the algebra $\mathfrak{R}$ of all classes $\mathfrak{s}(b)$ under the correspondence $b\longleftrightarrow\mathfrak{s}(b)$ in accordance with the properties*

(1) $\mathfrak{s}(b)=\mathfrak{s}(c)$ *if and only if* $b=c$;
(2) $\mathfrak{s}(b+c)=\mathfrak{s}(b)\Delta\mathfrak{s}(c)$;
(3) $\mathfrak{s}(bc)=\mathfrak{s}(b)\mathfrak{s}(c)$;
(4) $\mathfrak{s}(b\vee c)=\mathfrak{s}(b)\cup\mathfrak{s}(c)$.†

If $b=c$, then $ab=ac$ for every a in $\mathfrak{s}$, and hence $\mathfrak{s}(b)=\mathfrak{s}(c)$. On the other hand if $\mathfrak{s}(b)=\mathfrak{s}(c)$, then $ab=ac$ for every a in $\mathfrak{s}$ by virtue of Definition 3; hence $ab+ac=0$, $a(b+c)=0$, $b+c=0$, $b=c$, in accordance with Definition 5. In order that the element a in $\mathfrak{s}$ belong to $\mathfrak{s}(b+c)$ it is necessary and sufficient that $ab+ac=a(b+c)\neq 0$; since $ab+ac=0$ whenever $ab=ac$, we see that the relation $a(b+c)\neq 0$ holds if and only if one, but not both, of the pairs of relations $ab=a$, $ac=0$ and $ab=0$, $ac=a$ is valid; and we therefore conclude that property (2) holds. In similar fashion, we establish properties (3) and (4). We may also deduce (4) from (2) and (3), as follows:

$$\begin{aligned}\mathfrak{s}(b\vee c) &= \mathfrak{s}(b+c+bc)=\mathfrak{s}(b)\Delta\mathfrak{s}(c)\Delta\mathfrak{s}(bc)\\ &= \mathfrak{s}(b)\Delta\mathfrak{s}(c)\Delta\mathfrak{s}(b)\mathfrak{s}(c)=\mathfrak{s}(b)\cup\mathfrak{s}(c).\end{aligned}$$

THEOREM 9. *An atomic basis $\mathfrak{s}$ is a complete atomic system.*

We must show that there is no element b such that $b\neq 0$, $ba=0$ for every element a in $\mathfrak{s}$. Now, if there were, we should have $b=a_1+\cdots+a_n$ for some

† Here, as elsewhere in this paper, we use the symbols $\cup$ and Δ to designate union and union (modulo 2), or symmetric difference, for classes; and we indicate the formation of the intersection by juxtaposition of the symbols for the classes affected.

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

elements $a_1, \cdots, a_n$ in $\mathfrak{s}$, by Definition 4; we may suppose that these elements are distinct because of the law $a+a=0$. Thus we should obtain the relations $a_1=a_1(a_1+\cdots+a_n)=a_1b=0$ in accordance with Theorem 6 but in contradiction to the fact that $a_1\neq 0$.

THEOREM 10. *The representation of an element b as the sum of elements in a complete atomic system $\mathfrak{s}$ is unique: the summands are precisely the elements of the class $\mathfrak{s}(b)$, $b\neq 0$.*

If $b=a_1+\cdots+a_n$, the elements $a_1, \cdots, a_n$ in $\mathfrak{s}$ being taken as distinct, and if a is an arbitrary element in $\mathfrak{s}$, then $ab=aa_1+\cdots+aa_n$ is different from 0 if and only if one of the relations $a=a_1, \cdots, a=a_n$ is valid, as we see by virtue of Theorem 6. Hence $a_1, \cdots, a_n$ are precisely the elements of the class $\mathfrak{s}(b)$, $b\neq 0$, described in Theorem 8.

From Theorems 8, 9, and 10 we now obtain the following result:

THEOREM 11. *In order that a Boolean ring A contain an atomic basis $\mathfrak{s}$, it is necessary and sufficient that A be isomorphic to the algebra of all finite subclasses of a fixed finite or infinite class Σ, the elements in $\mathfrak{s}$ being in one-to-one correspondence with those of Σ. In particular such a ring A has a unit if and only if the classes $\mathfrak{s}$ and Σ are finite.*

When A has an atomic basis, the theorem follows at once from the preceding results, as we have already indicated. On the other hand, when A is isomorphic to an algebra of classes of the type described, it is evident that the elements of A corresponding to the one-element subclasses of Σ constitute an atomic basis in A. If a Boolean ring A containing an atomic basis $\mathfrak{s}$ has a unit e, then it is evident that $\mathfrak{s}=\mathfrak{s}(e)$ and hence that $\mathfrak{s}$ and the corresponding class Σ must both be finite. The converse result is also obvious.

THEOREM 12. *A finite Boolean ring with at least two elements contains an atomic basis $\mathfrak{s}$ and is therefore isomorphic to the algebra of all subclasses of a finite class Σ in one-to-one correspondence with $\mathfrak{s}$.*

In view of Theorem 11, it is sufficient for us to show that the given Boolean ring A has an atomic basis $\mathfrak{s}$. First we shall show that, if $a\neq 0$ is an arbitrary element in A, then there exists an atomic element contained in a. If a_1 is not an atomic element, there exists an element a such that $a_1a\neq 0$, $a_1a\neq a_1$. We denote a_1a by a_2 and observe that $a_1>a_2$. Let us suppose that there exist in A distinct non-zero elements $a_1, \cdots, a_k$ such that $a_1>a_2>\cdots>a_{k-1}>a_k$. If a_k is not an atomic element, then there exists an element a such that $a_ka\neq 0$, $a_ka\neq a_k$. We denote a_ka by a_{k+1} and observe that $a_k>a_{k+1}$. Thus we see, in view of the finiteness of A, that by virtue of this inductive construction there exist an integer n and elements $a_1, \cdots, a_n$,

where $a_1 > \cdots > a_n$ and a_n is an atomic element. We next designate by $\mathfrak{s}$ the class of all atomic elements in A and prove that $\mathfrak{s}$ is an atomic basis. Since $\mathfrak{s}$ is finite, its elements have a sum a. If b is an arbitrary element in A, we form the element $b+ab$ and apply the result just established: unless $b+ab=0$, there exists an atomic element c contained in $b+ab$. Now such an element c obviously has the properties $ac=c$, $(b+ab)c=c$; but these properties imply that $c=(b+ab)c=bc+bac=bc+bc=0$. We conclude therefore that $b+ab=0$ or, equivalently, $ab=b$. It follows that a is the unit e of the Boolean ring A and that any non-zero element b is expressible as the sum of atomic elements in $\mathfrak{s}$ through the equation $b=ab$. We have thereby shown that $\mathfrak{s}$ is an atomic basis.

We may observe that this theorem yields a new proof of the last part of Theorem 1, according to which every finite Boolean ring has a unit and has a cardinal number of the form 2^M. It shows further that when $M \geqq 1$ this integer is the number of elements in the atomic basis $\mathfrak{s}$.

In order to complete the consideration of finite Boolean rings, we state the following theorem without formal proof:

THEOREM 13. *A finite Boolean ring with exactly one element is isomorphic to the algebra consisting of the void class.*

CHAPTER II. SUBRINGS, IDEALS, AND HOMOMORPHISMS

1. **Subrings and their combinations.** A non-void subclass of an algebraic system is called a subsystem if it is closed under the fundamental operations of the system; that is, if the application of these operations to elements of the subclass yields elements of that class. The subsystems of a Boolean ring A are thus the subclasses of A which contain $a+b$ and ab whenever they contain a and b. The relation $a-b=a+b$, which is an immediate consequence of Theorem 1, therefore serves to identify the subsystems of A with the subrings of A.† Furthermore, the fact that the law of idempotence holds in any subclass of A shows that the subrings of A are Boolean rings in the sense of Definition 1. A few simple properties of the subrings of a Boolean ring A may be noted here without formal proof. Thus a subring $\mathfrak{a}$ has a unit a if and only if it contains an element a such that $ab=b$ or, equivalently, $a>b$ for every element b in $\mathfrak{a}$. Every subring contains the element 0 of A as its zero element. A subclass of a subring $\mathfrak{a}$ is a subring of A if and only if it is a subring of $\mathfrak{a}$. We may also cite the following examples of subrings: the subclass $\mathfrak{o}$ consisting of the element 0 alone; the subclass $\mathfrak{e}$ consisting of all the elements of A; and the subclass $\mathfrak{a}(a)$ consisting of all the elements b such that $ab=b$

† B. L. van der Waerden, *Moderne Algebra*, Berlin, 1930, vol. I, p. 53.

or, equivalently, $b<a$. We observe that $\mathfrak{o}=\mathfrak{a}(0)$; and that $\mathfrak{e}=\mathfrak{a}(e)$ in the case where A has a unit e.

Before passing to the study of combinations of subrings, we pause to consider briefly those non-void subclasses of a Boolean ring which are closed under the operations $\vee$ and $\cdot$ of Theorems 2–4. In view of the relation $a \vee b = a+b+ab$, every subring is such a subclass; but there exist subclasses with the indicated property which are not subrings, as is shown by the example of the subclass consisting of the element a alone, where $a \neq 0$, and by the example of the subclass consisting of the elements 0, a, b, ab, $a \vee b$, where a, b, and 0 are distinct (this subclass consists of three elements when $ab=a$ or when $ab=b$ and of five elements otherwise; Theorem 1 thus shows that it is not a subring). It is easily verified that these subclasses are instances of those algebraic systems variously known as C-lattices, distributive lattices, or arithmetic structures.† More than this is true: for recently MacNeille has proved that every distributive lattice is contained in a Boolean algebra (Boolean ring with unit) as a subclass of the indicated type by virtue of a strictly algebraic construction, of which the imbedding process given in the proof of Theorem 1 above is a special and very much simplified instance.‡

In any algebraic system, the intersection of any class of subsystems is itself a subsystem except, of course, in the case where it is void;§ the intersection is obviously the greatest subsystem contained in all subsystems of the given class. Hence any non-void subclass of an algebraic system generates a least subsystem containing it, namely, the intersection of all subsystems containing it. The subsystem thus generated by a given non-void subclass may be characterized alternatively as the class of all elements which can be constructed as "polynomials" in terms of the elements of the given subclass and of the fundamental operations of the system. It is understood in this statement that certain of the fundamental operations may, in a single application, affect infinitely many elements of the system. If $\mathfrak{A}$ is a non-void class of subsystems $\mathfrak{a}$, we may therefore define the sum and product of the subsystems in $\mathfrak{A}$ as follows: the sum, denoted by $S_{\mathfrak{a}\in\mathfrak{A}}\mathfrak{a}$, is the least subsystem containing every $\mathfrak{a}$ in $\mathfrak{A}$, or, equivalently, is the subsystem generated by the union

† The term C-lattice is used by Garrett Birkhoff, Proceedings of the Cambridge Philosophical Society, vol. 29 (1933), pp. 441–464; distributive lattice by MacNeille, Harvard doctoral dissertation, *The Theory of Partially Ordered Sets*, 1935; arithmetic structure by Ore, Annals of Mathematics, (2), vol. 36 (1935), pp. 406–437.

‡ MacNeille, doctoral dissertation, *The Theory of Partially Ordered Sets*, 1935, not yet published. A summary is given in the Proceedings of the National Academy of Sciences, vol. 22 (1936), pp. 45–50.

§ For some purposes it is convenient to regard the void class as a subsystem, but that is not the case in the present paper.

$\Sigma_{\alpha\in\mathfrak{A}}\mathfrak{a}$ of the classes $\mathfrak{a}$ belonging to $\mathfrak{A}$; and the product, denoted by $P_{\alpha\in\mathfrak{A}}\mathfrak{a}$, is the intersection $\Pi_{\alpha\in\mathfrak{A}}\mathfrak{a}$ of the classes $\mathfrak{a}$ in $\mathfrak{A}$, when it is not void. In the special case where $\mathfrak{A}$ consists of two subsystems $\mathfrak{a}$ and $\mathfrak{b}$, we write the sum as $\mathfrak{a}\vee\mathfrak{b}$ and the product as $\mathfrak{a}\cdot\mathfrak{b}$ or simply $\mathfrak{ab}$. In general, the product of subsystems may fail to exist; but in the case of Boolean rings this difficulty is removed by the fact that the element 0 is common to all subrings. It is important for us to examine the specialization of these general concepts to the case of Boolean rings.

In the first place we shall give more detailed information concerning the subring generated by a given subclass. We have

THEOREM 14. *In order that the subring* $\mathfrak{a}(\mathfrak{s})$ *generated by a non-void subclass* $\mathfrak{s}$ *of a Boolean ring A possess a unit, it is necessary and sufficient that* $\mathfrak{s}$ *contain elements* $a_1, \cdots, a_n$ *such that* $b<a_1\vee\cdots\vee a_n$ *for every element* b *in* $\mathfrak{s}$. *When this condition is satisfied, the element* $a=a_1\vee\cdots\vee a_n$ *is the unit of* $\mathfrak{a}(\mathfrak{s})$; *and* $\mathfrak{a}(\mathfrak{s})$ *is the class of all elements which can be constructed as polynomials in terms of elements* b *and* $a+b$, *where* b *is in* $\mathfrak{s}$, *and of the operations* $\vee$ *and* $\cdot$ *alone. In particular, if A has a unit e and* $\mathfrak{s}$ *contains e, then* $\mathfrak{a}(\mathfrak{s})$ *is the class of all elements which can be constructed as polynomials in terms of elements* b *and* $b'=b+e$, *where* b *is in* $\mathfrak{s}$, *and of the operations* $\vee$ *and* $\cdot$ *alone.*

The existence of elements $a_1, \cdots, a_n$ in $\mathfrak{s}$ such that $a_1\vee\cdots\vee a_n>b$ for every element b in $\mathfrak{s}$ leads immediately to the conclusion that $a=a_1\vee\cdots\vee a_n$ is in $\mathfrak{a}(\mathfrak{s})$ and is its unit: for, if c is any element of $\mathfrak{a}(\mathfrak{s})$, then $c=q(b_1, \cdots, b_m)$ where q is a polynomial in terms of elements $b_1, \cdots, b_m$ in $\mathfrak{s}$ and of the operations $+$ and $\cdot$ of the Boolean ring; and, by virtue of the distributive law and the law of idempotence, such an element c satisfies the relations $ac=aq(b_1, \cdots, b_m)=q(ab_1, \cdots, ab_m)=q(b_1, \cdots, b_m)=c$. On the other hand, if $\mathfrak{a}(\mathfrak{s})$ has a unit a, we can express a as a polynomial $r(a_1, \cdots, a_n)$, where $a_1, \cdots, a_n$ are in $\mathfrak{s}$, and have to show, in order to complete the discussion, that a can be expressed as $a_1\vee\cdots\vee a_n$. Since $a_1\vee\cdots\vee a_n$ is in $\mathfrak{a}(\mathfrak{s})$ and since a is the unit in $\mathfrak{a}(\mathfrak{s})$, we have $a_1\vee\cdots\vee a_n<a$. On applying the general relations $b\vee c>b+c$, $b>bc$, $c>bc$ to the polynomial $r(a_1, \cdots, a_n)$ we find that $a_1\vee\cdots\vee a_n>a$. Hence it is true that $a=a_1\vee\cdots\vee a_n$, as we wished to show. In a subring $\mathfrak{a}(\mathfrak{s})$ with unit a, we can apply the relations (1), (2), (6), and (7) of Theorem 2 with appropriate change of letters. Thus if $c=q(b_1, \cdots, b_m)$ where $b_1, \cdots, b_m$ are in $\mathfrak{s}$ we can use these relations to write $q(b_1, \cdots, b_m)=q^*(b_1, \cdots, b_m, b_1', \cdots, b_m')$ where q^* is a polynomial in terms of the operations $\vee$ and $\cdot$ alone, and $b_k'=b_k+a$ for $k=1, \cdots, m$. The final assertions of the present theorem follow at once from this result.

Using the notations for sums and products of subrings introduced above

together with the customary symbol $\subset$ for the relation of class-inclusion, we now state the principle facts concerning the operations upon subsystems. While we phrase them in terms of Boolean rings, they are easily seen to be valid in quite arbitrary algebraic systems.

THEOREM 15. *If A is a Boolean ring, then the class $\mathfrak{A}$ of all subrings of A has the following properties under the operations of addition and multiplication introduced above*:

(1) $\mathfrak{a} \vee \mathfrak{b} = \mathfrak{b} \vee \mathfrak{a}$; (2) $\mathfrak{ab} = \mathfrak{ba}$;

(3) $\mathfrak{a} \vee (\mathfrak{b} \vee \mathfrak{c}) = (\mathfrak{a} \vee \mathfrak{b}) \vee \mathfrak{c}$; (4) $\mathfrak{a}(\mathfrak{bc}) = (\mathfrak{ab})\mathfrak{c}$;

(5) $\mathfrak{a}(\mathfrak{b} \vee \mathfrak{c}) \supset \mathfrak{ab} \vee \mathfrak{ac}$; (6) $(\mathfrak{a} \vee \mathfrak{b})(\mathfrak{a} \vee \mathfrak{c}) \supset \mathfrak{a} \vee \mathfrak{bc}$;

(7) $\mathfrak{a} \vee \mathfrak{a} = \mathfrak{a}$; (8) $\mathfrak{aa} = \mathfrak{a}$;

(9) $\mathfrak{a} \subset \mathfrak{b}$ *if and only if* $\mathfrak{ab} = \mathfrak{a}$;

(10) *if $\mathfrak{B}$ is a non-void class of non-void classes $\mathfrak{B}$ of subrings $\mathfrak{a}$ of A, and if $\mathfrak{C}$ is the union $\Sigma\mathfrak{B}_{\mathfrak{B}\in\mathfrak{B}}$, then*

$$\mathop{S}_{\mathfrak{B}\in\mathfrak{B}} \left(\mathop{S}_{\mathfrak{a}\in\mathfrak{B}} \mathfrak{a}\right) = \mathop{S}_{\mathfrak{a}\in\mathfrak{C}} \mathfrak{a};$$

(11) *if $\mathfrak{B}$, $\mathfrak{B}$, and $\mathfrak{C}$ have the same significance as in* (10), *then*

$$\mathop{P}_{\mathfrak{B}\in\mathfrak{B}} \left(\mathop{P}_{\mathfrak{a}\in\mathfrak{B}} \mathfrak{a}\right) = \mathop{P}_{\mathfrak{a}\in\mathfrak{C}} \mathfrak{a};$$

(12) *if $\mathfrak{b}$ is any subring of A and $\mathfrak{B}$ is any non-void class of subrings $\mathfrak{a}$ of A, then*

$$\mathfrak{b}\left(\mathop{S}_{\mathfrak{a}\in\mathfrak{B}} \mathfrak{a}\right) \supset \mathop{S}_{\mathfrak{a}\in\mathfrak{B}} (\mathfrak{ba});$$

(13) *if $\mathfrak{b}$ and $\mathfrak{B}$ have the same significance as in* (12), *then*

$$\mathop{P}_{\mathfrak{a}\in\mathfrak{B}} (\mathfrak{b} \vee \mathfrak{a}) \supset \mathfrak{b} \vee \mathop{P}_{\mathfrak{a}\in\mathfrak{B}} \mathfrak{a}.$$

The special subrings $\mathfrak{o}$ and $\mathfrak{e}$ have the following properties:

(14) $\mathfrak{oa} = \mathfrak{o}$, $\mathfrak{a} \vee \mathfrak{o} = \mathfrak{a}$;

(15) $\mathfrak{ea} = \mathfrak{a}$, $\mathfrak{a} \vee \mathfrak{e} = \mathfrak{e}$.

Except for properties (5), (6), (12), and (13), all these properties are easily verified by quite trivial arguments; and properties (5) and (6) are special cases of (12) and (13) respectively. We shall therefore confine our discussion to the two latter properties. To establish (12) we proceed as follows: an element belongs to the subring on the right of (12) if and only if it is a polynomial in terms of elements simultaneously in the subring $\mathfrak{b}$ and in the sub-

rings $\mathfrak{a}$; on the other hand, an element belongs to the subring on the left of (12) if and only if it is simultaneously a polynomial in terms of elements in $\mathfrak{b}$ and a polynomial in terms of elements in the subrings $\mathfrak{a}$; from these algebraic descriptions, the inclusion-relation (12) is evident. We prove (13) in a similar manner: an element belongs to the subring on the right of (13) if and only if it is a polynomial in terms of elements in the subring $\mathfrak{b}$ and common to all the subrings $\mathfrak{a}$; and an element belongs to the subring on the left of (13) if and only if it is, simultaneously for all subrings $\mathfrak{a}$, a polynomial in terms of elements of the subring $\mathfrak{b}$ and of a subring $\mathfrak{a}$. It is of interest to show that the relations (5) and (6), and hence also the relations (12) and (13), cannot be strengthened. In a Boolean ring with unit and with four or more elements, let a be an element distinct from 0 and from e; let $\mathfrak{a}$ be the subring consisting of the elements 0, a; let $\mathfrak{b}$ be the subring consisting of the elements 0, e; let $\mathfrak{c}$ be the subring consisting of the elements 0, a', where $a'=a+e$; and let $\mathfrak{d}$ be the subring consisting of the elements 0, a, a', e. We then see that $\mathfrak{a}(\mathfrak{b}\vee\mathfrak{c})=\mathfrak{a}$, $\mathfrak{a}\mathfrak{b}\vee\mathfrak{a}\mathfrak{c}=\mathfrak{o}$, $(\mathfrak{a}\vee\mathfrak{b})(\mathfrak{a}\vee\mathfrak{c})=\mathfrak{d}$, $\mathfrak{a}\vee\mathfrak{b}\mathfrak{c}=\mathfrak{a}$, $\mathfrak{o}\neq\mathfrak{a}\neq\mathfrak{d}$.

2. **Ideals and their combinations.** We shall now turn to the study of those special subrings known as invariant subrings or ideals, characterized by the property of containing a and ab together whatever the element b. Thus a non-void subclass of a Boolean ring is an ideal if and only if it contains $a+b$ together with a and b, and c together with a whenever $c<a$. Since ideals are special subrings, the discussion of the preceding section applies to them at once; but, as we shall see below, is capable of being made considerably more precise by virtue of the restriction to ideals. We shall note a few simple properties of ideals without formal proof. Thus, an ideal $\mathfrak{a}$ in a Boolean ring A has a unit a if and only if it consists of all elements c such that $c<a$; in other words, if and only if it is identical with the subring $\mathfrak{a}(a)$ introduced above and seen now to be an ideal. Moreover an ideal $\mathfrak{a}$ which contains the element a necessarily contains the ideal $\mathfrak{a}(a)$. A subclass $\mathfrak{b}$ of an ideal $\mathfrak{a}$ in a Boolean ring A is an ideal in $\mathfrak{a}$, regarded as a Boolean ring, if and only if it is an ideal in A. If $\mathfrak{a}$ is an ideal and $\mathfrak{b}$ a subring, then $\mathfrak{a}$ is an ideal in the subring $\mathfrak{a}\vee\mathfrak{b}$ and $\mathfrak{a}\mathfrak{b}$ is an ideal in the subring $\mathfrak{b}$. The subrings $\mathfrak{o}$, $\mathfrak{a}(a)$, $\mathfrak{e}$, previously introduced, are all ideals. It is important for us to recall the arithmetical terminology used to describe the inclusion relation $\mathfrak{a}\subset\mathfrak{b}$ between ideals: if $\mathfrak{a}$ is contained in $\mathfrak{b}$, then $\mathfrak{b}$ is said to divide $\mathfrak{a}$ or to be a divisor of $\mathfrak{a}$, and $\mathfrak{a}$ is said to be divisible by $\mathfrak{b}$. Thus the product $\mathfrak{a}\mathfrak{b}$ of ideals $\mathfrak{a}$ and $\mathfrak{b}$ is divisible by its factors $\mathfrak{a}$ and $\mathfrak{b}$, the product being itself an ideal in accordance with a result established below.

The distinction which it was necessary to draw between subsystems of a Boolean ring defined in terms of the operations $+$ and $\cdot$ and those defined in

terms of $\vee$ and $\cdot$ vanishes in the case of multiplicatively invariant subsystems, as indicated in the following theorem:

THEOREM 16. *In order that a non-void subclass* $\mathfrak{a}$ *of a Boolean ring* A *be an ideal it is necessary and sufficient that*

(1) $\mathfrak{a}$ *contain* $a \vee b$ *together with* a *and* b,

(2) $\mathfrak{a}$ *contain* ab *whenever it contains* a;

or, equivalently, that

(1) $\mathfrak{a}$ *contain* $a \vee b$ *together with* a *and* b,

(2′) $\mathfrak{a}$ *contain* c *together with* a *whenever* $c < a$.

Since the conditions (2) and (2′) are equivalent, we need consider only conditions (1) and (2). The necessity of the latter conditions is evident. To establish their sufficiency, it is enough to show that they imply that $\mathfrak{a}$ contains $a+b$ together with a and b. Now if a and b are in $\mathfrak{a}$, so are $a \vee b$ and $(a \vee b)(a+b)$ by virtue of (1) and (2); but the relation

$$(a \vee b)(a + b) = [(a + b) + ab](a + b) = (a + b) + (ab + ab) = a + b$$

shows that $a+b$ is also in $\mathfrak{a}$, as we wished to prove.

Before considering the specialization of Theorem 15 to the case of ideals, it is necessary for us to indicate some particular results analogous to those presented in Theorem 14.

THEOREM 17. *If* $\mathfrak{s}$ *is an arbitrary non-void subclass of a Boolean ring* A *and if* $\mathfrak{a}(\mathfrak{s})$ *is the class of all elements* a *such that* $a < a_1 \vee \cdots \vee a_n$ *for appropriate elements* $a_1, \cdots, a_n$ *in* $\mathfrak{s}$, *then* $\mathfrak{a}(\mathfrak{s})$ *is an ideal; and every ideal containing* $\mathfrak{s}$ *contains* $\mathfrak{a}(\mathfrak{s})$. *The ideal* $\mathfrak{a}(\mathfrak{s})$ *may be characterized alternatively as the class of all elements* a *such that* $a = a_1b_1 \vee \cdots \vee a_nb_n$ *where* $a_1, \cdots, a_n$ *are in* $\mathfrak{s}$ *and* $b_1, \cdots, b_n$ *are in* A. *If* $\mathfrak{s}$ *is the union of the ideals* $\mathfrak{a}$ *in a given class* $\mathfrak{B}$, *then* $\mathfrak{a}(\mathfrak{s})$ *is the class of all elements* a *such that* $a = a_1 \vee \cdots \vee a_n$ *where* a_k *is in* $\mathfrak{a}_k$ *and* $\mathfrak{a}_k$ *in* $\mathfrak{B}$ *for* $k = 1, \cdots, n$.

From Theorem 16, it is evident that an ideal containing the class $\mathfrak{s}$ must contain every element $a_1 \vee \cdots \vee a_n$ where $a_1, \cdots, a_n$ are in $\mathfrak{s}$, and hence every element a such that $a < a_1 \vee \cdots \vee a_n$. Thus $\mathfrak{a}(\mathfrak{s})$ is contained in every ideal which contains $\mathfrak{s}$. To show that $\mathfrak{a}(\mathfrak{s})$ is itself an ideal, we appeal again to Theorem 16. It is evident that condition (2′) of that theorem is satisfied in the present instance. It is easily verified that condition (1) also holds: for, if $a < a_1 \vee \cdots \vee a_n$ and $b < b_1 \vee \cdots \vee b_p$ where $a_1, \cdots, a_n, b_1, \cdots, b_p$ are in $\mathfrak{s}$, then $a \vee b < a_1 \vee \cdots \vee a_n \vee b_1 \vee \cdots \vee b_p$. The equivalent characterization of $\mathfrak{a}(\mathfrak{s})$ is established as follows: if $a < a_1 \vee \cdots \vee a_n$, then $a = a_1b_1 \vee \cdots \vee a_nb_n$ with $b_1 = \cdots = b_n = a$; and, if $a = a_1b_1 \vee \cdots \vee a_nb_n$, then $a < a_1 \vee \cdots \vee a_n$. Finally, to establish the characterization of $\mathfrak{a}(\mathfrak{s})$ when $\mathfrak{s}$ is the union of ideals $\mathfrak{a}$,

we first express the element a in $\mathfrak{a}(\mathfrak{s})$ in the form $a = a_1 b_1 \vee \cdots \vee a_n b_n$ where $a_1, \cdots, a_n$ are in $\mathfrak{s}$. We then observe that there must exist ideals $\mathfrak{a}_1, \cdots, \mathfrak{a}_n$ of the given class $\mathfrak{B}$ which contain $a_1, \cdots, a_n$ respectively. It follows that $a_k b_k$ is an element of the ideal $\mathfrak{a}_k$. Thus an obvious change of notation permits us to write $a = a_1 \vee \cdots \vee a_n$ where a_k is in $\mathfrak{a}_k$ and $\mathfrak{a}_k$ in $\mathfrak{B}$, for $k = 1, \cdots, n$. Conversely, every such element is in the ideal $\mathfrak{a}(\mathfrak{s})$, as we have already seen.

With the help of Theorem 17, we can now obtain the desired counterpart of Theorem 15.

THEOREM 18. *In a Boolean ring A, the subrings obtained as sums or as products of ideals are themselves ideals; in other words, the class $\mathfrak{J}$ of all ideals in A is a subsystem of the system $\mathfrak{A}$ of all subrings of A under the unrestricted operations of addition and multiplication. The properties of these operations which hold in $\mathfrak{A}$ hold also in $\mathfrak{J}$, with the refinement that properties* (5), (6), *and* (12) *of Theorem* 15 *are to be replaced respectively by the following sharper properties, to which we give the corresponding numbers*:

(5) $\mathfrak{a}(\mathfrak{b} \vee \mathfrak{c}) = \mathfrak{a}\mathfrak{b} \vee \mathfrak{a}\mathfrak{c}$; (6) $(\mathfrak{a} \vee \mathfrak{b})(\mathfrak{a} \vee \mathfrak{c}) = \mathfrak{a} \vee \mathfrak{b}\mathfrak{c}$;

(12) *if $\mathfrak{b}$ is an ideal and $\mathfrak{B}$ a non-void class of ideals $\mathfrak{a}$, then*

$$\mathfrak{b} \sum_{\mathfrak{a} \in \mathfrak{B}} \mathfrak{a} = \sum_{\mathfrak{a} \in \mathfrak{B}} \mathfrak{b}\mathfrak{a}.$$

The ideal $\mathfrak{a}\mathfrak{b}$, where $\mathfrak{a}$ and $\mathfrak{b}$ are ideals, is the class of elements c where $c = ab$, a in $\mathfrak{a}$ and b in $\mathfrak{b}$. The ideal $\mathfrak{a}(\mathfrak{s})$ of Theorem 17 *is the product of all ideals containing $\mathfrak{s}$; and, in the particular case where $\mathfrak{s}$ is the union of a class of ideals, $\mathfrak{a}(\mathfrak{s})$ is the sum of the ideals in that class.*

If a is any element in the intersection of ideals $\mathfrak{a}$, then ab is also in their intersection whatever the element b. Thus the subring which is the product of the ideals $\mathfrak{a}$ is an ideal. In particular the product $\mathfrak{a}\mathfrak{b}$ of ideals $\mathfrak{a}$ and $\mathfrak{b}$ is an ideal and consists of those elements c such that $c = ab$ where a is in $\mathfrak{a}$ and b in $\mathfrak{b}$: for any such element is common to $\mathfrak{a}$ and $\mathfrak{b}$; and, if c is common to $\mathfrak{a}$ and $\mathfrak{b}$, then $c = ab$ where $a = c$ and $b = c$. Theorem 17 now shows that the ideal $\mathfrak{a}(\mathfrak{s})$ is the product, or, equivalently, the intersection, of all the ideals containing $\mathfrak{s}$, the ideals $\mathfrak{e}$ and $\mathfrak{a}(\mathfrak{s})$ both having the latter property. Now, if $\mathfrak{s}$ is the union of ideals $\mathfrak{a}$, we see from Theorem 17 that $\mathfrak{a}(\mathfrak{s})$ is contained in the subring which is the sum of the ideals $\mathfrak{a}$; but, since $\mathfrak{a}(\mathfrak{s})$ is a subring containing $\mathfrak{s}$, it must contain also the sum in question. It follows that any sum of ideals is identical with the ideal $\mathfrak{a}(\mathfrak{s})$, where $\mathfrak{s}$ is the union of those ideals. The preceding results evidently serve to establish the assertion that $\mathfrak{J}$ is a subsystem of $\mathfrak{A}$. It remains for us to establish the sharper forms of (5), (6), and (12). Since (12) implies (5), we confine our discussion to (6) and (12). In view of Theorem 15,

it is sufficient for us to show that $\mathfrak{b}S_{a\mathfrak{B}}\mathfrak{a} \subset S_{a\mathfrak{B}}\mathfrak{ba}$, $(\mathfrak{a} \vee \mathfrak{b})(\mathfrak{a} \vee \mathfrak{c}) \subset \mathfrak{a} \vee \mathfrak{bc}$. Using the results of Theorem 17 together with those just proved, we see that every element in $\mathfrak{b}S_{a\mathfrak{B}}\mathfrak{a}$ is expressible in the form $b(a_1 \vee \cdots \vee a_n) = ba_1 \vee \cdots \vee ba_n$ where b is in $\mathfrak{b}$ and a_k in an ideal $\mathfrak{a}_k$ of the class $\mathfrak{B}$. Since ba_k is in the ideal $\mathfrak{ba}_k$, we conclude that every element of the indicated form is in the ideal $S_{a\mathfrak{B}}\mathfrak{ba}$, thus establishing (12). Similarly, we see that every element in $(\mathfrak{a} \vee \mathfrak{b})(\mathfrak{a} \vee \mathfrak{c})$ can be expressed in the form $(a_1 \vee b)(a_2 \vee c)$ where a_1 and a_2 are in $\mathfrak{a}$, b is in $\mathfrak{b}$, and c is in $\mathfrak{c}$; and, since $(a_1 \vee b)(a_2 \vee c) = (a_1a_2 \vee a_1c \vee a_2b) \vee bc$ where $a_1a_2 \vee a_1c \vee a_2b$ is in $\mathfrak{a}$ and bc is in $\mathfrak{bc}$, we see that every such element is in $\mathfrak{a} \vee \mathfrak{bc}$, thus establishing (6).

It is of interest to remark that property (13) of Theorem 15 cannot be similarly sharpened. This we shall show by examples to be given in a later paper. It is also of interest to remark that in the case of a general abstract ring the properties (5), (6), and (12) of Theorem 15 cannot be replaced by the sharper ones which have just been established in the case of Boolean rings. We note that under the finite operations, namely, the operations of forming the finite sum $\mathfrak{a} \vee \mathfrak{b}$ and the finite product $\mathfrak{ab}$, the system $\mathfrak{J}$ is a distributive lattice by virtue of the sharpened properties (5) and (6). The Boolean rings are thus special instances of those rings in which the ideals constitute a distributive lattice.†

We shall now introduce in the class $\mathfrak{J}$ of ideals a unary operation in many respects analogous to the operation $'$ defined in a Boolean ring with unit. Such an operation can be defined in any commutative ring and can be suitably generalized even in the case of a non-commutative ring. We shall investigate its properties only in the case immediately before us. It will be helpful to make use of the following terminology:

DEFINITION 6. *Two elements a and b in a Boolean ring are said to be orthogonal if $ab=0$; and two non-void subclasses of a Boolean ring are said to be orthogonal if every element of one is orthogonal to every element of the other.*

As a basis for our definition of the desired operation, we first establish the following result:

THEOREM 19. *If $\mathfrak{s}$ is any non-void subclass of a Boolean ring A, then the class $\mathfrak{s}'$ of all elements orthogonal to every element of $\mathfrak{s}$ is an ideal in A which is orthogonal to $\mathfrak{s}$ and contains every subclass of A orthogonal to $\mathfrak{s}$. Two ideals $\mathfrak{a}$ and $\mathfrak{b}$ are orthogonal if and only if $\mathfrak{ab}=\mathfrak{o}$.*

It is obvious that $\mathfrak{s}'$ contains the element 0, that it is orthogonal to $\mathfrak{s}$, and that it contains every subclass of A orthogonal to $\mathfrak{s}$. If a and b are in $\mathfrak{s}'$,

† The indicated class of rings has been discussed by Garrett Birkhoff, under certain strong restrictions, Bulletin of the American Mathematical Society, vol. 40 (1934), pp. 613–619.

then $a+b$ is in $\mathfrak{s}'$: for we have $(a+b)c=ac+bc=0+0=0$ for every element c in $\mathfrak{s}$. Similarly, if a is in $\mathfrak{s}'$ and b is in A, then ab is in $\mathfrak{s}'$: for we have $(ab)c=(ac)b=0b=0$ for every element c in $\mathfrak{s}$. We see therefore that $\mathfrak{s}'$ is an ideal. The characterization of the ideal product $\mathfrak{ab}$ given in Theorem 18 shows at once that $\mathfrak{a}$ and $\mathfrak{b}$ are orthogonal if and only if $\mathfrak{ab}=\mathfrak{o}$.

We can now state our fundamental definition.

DEFINITION 7. *The ideal $\mathfrak{s}'$ associated with an arbitrary non-void subclass $\mathfrak{s}$ of a Boolean ring A in the manner indicated in Theorem* 19 *is called the orthogonal complement, or, more briefly, the orthocomplement of $\mathfrak{s}$; and the operation of forming the ideal $\mathfrak{s}'$ is called orthogonal complementation, or, more briefly, orthocomplementation. The orthocomplement of $\mathfrak{s}'$ is denoted by $\mathfrak{s}''$, that of $\mathfrak{s}''$ by $\mathfrak{s}'''$; and, more generally, the symbol $\mathfrak{s}^{(n)}$ is defined recursively for $n\geqq 1$ by the relations $\mathfrak{s}^{(1)}=\mathfrak{s}'$, $\mathfrak{s}^{(n+1)}=(\mathfrak{s}^{(n)})'$.*

The chief properties of the operation so defined are given in the three theorems which follow.

THEOREM 20. *The operation of orthocomplementation has the following general properties*:

(1) $\mathfrak{s}\subset\mathfrak{t}$ *implies* $\mathfrak{s}'\supset\mathfrak{t}'$;

(2) $\mathfrak{s}\subset\mathfrak{a}(\mathfrak{s})\subset\mathfrak{s}''$, *where* $\mathfrak{a}(\mathfrak{s})$ *is the ideal generated by* $\mathfrak{s}$;

(3) $\mathfrak{s}^{(m)}=\mathfrak{s}^{(n)}$ *when* m *and* n *are congruent* (mod 2), $\mathfrak{s}^{(m)}\mathfrak{s}^{(n)}=\mathfrak{o}$ *when* m *and* n *are not congruent* (mod 2); *in particular*, $\mathfrak{s}'''=\mathfrak{s}'$.

If $\mathfrak{s}$ is contained in $\mathfrak{t}$, then $\mathfrak{t}'$ is orthogonal to $\mathfrak{s}$ as well as to $\mathfrak{t}$ and must therefore be contained in $\mathfrak{s}'$ in accordance with Theorem 19. Since $\mathfrak{s}$ is orthogonal to $\mathfrak{s}'$ by definition, $\mathfrak{s}$ is contained in the ideal $\mathfrak{s}''$; and by Theorem 17 we have $\mathfrak{s}\subset\mathfrak{a}(\mathfrak{s})\subset\mathfrak{s}''$. The relation $\mathfrak{s}\subset\mathfrak{s}''$ implies $\mathfrak{s}'\supset\mathfrak{s}'''$; but we also have $\mathfrak{s}'''=(\mathfrak{s}')''\supset\mathfrak{s}'$ and therefore conclude that $\mathfrak{s}'''=\mathfrak{s}'$. By an obvious induction we now see that $\mathfrak{s}^{(n+2)}=\mathfrak{s}^{(n)}$, $\mathfrak{s}^{(n+1)}\mathfrak{s}^{(n)}=\mathfrak{o}$. A further induction leads to the general proposition (3).

THEOREM 21. *Within the class $\mathfrak{F}$ of all ideals in a Boolean ring A, the operation of orthocomplementation has the following special properties*:

(1) $\mathfrak{a}\subset\mathfrak{a}''$; (2) $\mathfrak{aa}'=\mathfrak{o}$; (3) $\mathfrak{o}'=\mathfrak{e}$, $\mathfrak{e}'=\mathfrak{o}$;

(4) *the orthocomplement of a sum is equal to the product of the orthocomplements of the summands; in particular*, $(\mathfrak{a}\vee\mathfrak{b})'=\mathfrak{a}'\mathfrak{b}'$;

(5) *the orthocomplement of a product contains the sum of the orthocomplements of the multiplicands; in particular*, $(\mathfrak{ab})'\supset\mathfrak{a}'\vee\mathfrak{b}'$.

The properties (1), (2), (3) follow at once from Theorems 19 and 20. To prove (4), we observed that, by virtue of Theorems 17 and 18, an element b is in $(S_{a\in\mathfrak{A}}\mathfrak{a})'$ if and only if $b(a_1\vee\cdots\vee a_n)=0$ for every set of elements

$a_1, \cdots, a_n$ belonging respectively to ideals $\mathfrak{a}_1, \cdots, \mathfrak{a}_n$ in the class $\mathfrak{B}$. Now $b(a_1 \vee \cdots \vee a_n) = ba_1 \vee \cdots \vee ba_n = 0$ if and only if $ba_1 = \cdots = ba_n = 0$, as is easily verified. The latter condition obviously holds for the indicated elements $a_1, \cdots, a_n$ if and only if b is in every ideal $\mathfrak{a}'$, where $\mathfrak{a}$ is in the class $\mathfrak{B}$. This result serves to establish (4). To prove (5), we observe that by virtue of Theorems 17 and 18 an element b in $S_{\mathfrak{a}\in\mathfrak{B}}\mathfrak{a}'$ is expressible in the form $b = a_1 \vee \cdots \vee a_n$ where a_k is in $\mathfrak{a}_k'$ and $\mathfrak{a}_k$ in $\mathfrak{B}$ for $k = 1, \cdots, n$; hence we see that, if c is an arbitrary element in the product of the ideals $\mathfrak{a}$, then $bc = (a_1 \vee \cdots \vee a_n)c = a_1c \vee \cdots \vee a_nc = 0$; and we conclude that b is in $(P_{\mathfrak{a}\in\mathfrak{B}}\mathfrak{a})'$, thus establishing (5).

The properties (1) and (5) of this theorem cannot be sharpened within the class $\mathfrak{I}$, save in the case of Boolean rings of very special type. We shall give relevant examples in another paper.

Since the product $\mathfrak{ab}$ of an ideal $\mathfrak{a}$ and a subring $\mathfrak{b}$ is an ideal in $\mathfrak{b}$, we may seek to determine the orthocomplement of $\mathfrak{ab}$ relative to $\mathfrak{b}$. This determination is possible when $\mathfrak{b}$ is an ideal, and proves to be important in subsequent developments.

THEOREM 22. *If* $\mathfrak{a}$ *and* $\mathfrak{b}$ *are ideals in a Boolean ring* A, *the orthocomplement* $\mathfrak{c}$ *of the ideal* $\mathfrak{ab}$ *in the subring* $\mathfrak{b}$ *satisfies the relation* $\mathfrak{c} = \mathfrak{a}'\mathfrak{b}$.

Since $\mathfrak{ab}$ and $\mathfrak{a}'\mathfrak{b}$ are ideals in $\mathfrak{b}$ such that $(\mathfrak{ab})(\mathfrak{a}'\mathfrak{b}) = (\mathfrak{aa}')\mathfrak{b} = \mathfrak{o}$, we see from Theorem 19 that $\mathfrak{c} \supset \mathfrak{a}'\mathfrak{b}$. On the other hand, we obviously have $\mathfrak{c} \subset \mathfrak{b}$, we can prove that $\mathfrak{c} \subset \mathfrak{a}'$, and we can therefore conclude that $\mathfrak{c} \subset \mathfrak{a}'\mathfrak{b}$ and hence $\mathfrak{c} = \mathfrak{a}'\mathfrak{b}$: for, if c is any element in $\mathfrak{c}$, the element ac is in $\mathfrak{ab}$ for every element a in $\mathfrak{a}$ and thus has the property $ac = (ac)c = 0$; but this property implies the relation $\mathfrak{c} \subset \mathfrak{a}'$.

In case $\mathfrak{b}$ is a subring but not an ideal, the relation $\mathfrak{c} \supset \mathfrak{a}'\mathfrak{b}$ is valid but cannot in general be replaced by the stronger relation $\mathfrak{c} = \mathfrak{a}'\mathfrak{b}$. To show this, we use the fact that in general the ideal $\mathfrak{a}$ can be chosen so that $\mathfrak{a} \vee \mathfrak{a}' \neq \mathfrak{e}$. If we then take b as an element not in $\mathfrak{a} \vee \mathfrak{a}'$ and let $\mathfrak{b}$ be the subring consisting of the elements 0, b, we see that $\mathfrak{ab} = \mathfrak{a}'\mathfrak{b} = \mathfrak{o}$, $\mathfrak{c} = \mathfrak{b} \neq \mathfrak{o}$.

3. **A classification of ideals.** The operation of orthocomplementation introduced at the end of the preceding section leads to an important classification of the ideals in a Boolean ring. While an exhaustive study of this classification would take us too far afield, some detailed knowledge of the behavior of ideals under the operation of orthocomplementation is essential in Chapter III. We shall therefore give only a partial investigation of this subject here, leaving for another paper the presentation of the complete theory. It is proper to point out in advance that the proposed classification degenerates only in a few very special types of Boolean ring, too simple to be of very great

interest in any other connection. We shall accordingly have no need to concern ourselves in the present section with the possibility of degeneracies. Our fundamental definition is the following:

DEFINITION 8. *In a Boolean ring A, an ideal $\mathfrak{a}$ is said to be*

(1) *principal if $\mathfrak{a}=\mathfrak{a}(a)$ for some element a;*

(2) *semiprincipal if $\mathfrak{a}=\mathfrak{a}(a)$ or $\mathfrak{a}=\mathfrak{a}'(a)$ for some element a;*

(3) *simple if $\mathfrak{a}\vee\mathfrak{a}'=\mathfrak{e}$;* (4) *normal if $\mathfrak{a}=\mathfrak{a}''$.*

The classes of principal, semiprincipal, simple, and normal ideals are denoted by the letters $\mathfrak{P}$, $\mathfrak{P}^*$, $\mathfrak{S}$, and $\mathfrak{N}$ respectively.

We may observe that the term "principal ideal" is here used with its ordinary significance: for Theorem 17 shows that the ideal $\mathfrak{a}(a)$ is the ideal generated by the class $\mathfrak{s}$ consisting of the element a alone.

The elementary relations between the classes of ideals thus introduced are given in the three following theorems.

THEOREM 23. *The classes defined in Definition* 8 *satisfy the following inclusion relations*:

(1) $$\mathfrak{P}\subset\mathfrak{P}^*\subset\mathfrak{S}\subset\mathfrak{N}\subset\mathfrak{J};$$

(2) $\mathfrak{P}$ *contains* $\mathfrak{o}$; (3) $\mathfrak{P}^*$ *contains* $\mathfrak{e}$.

The relations (2) and (3) are obvious. Of the inclusions in (1), the first and last are obvious. Hence we need discuss in detail only the relations $\mathfrak{P}^*\subset\mathfrak{S}$, $\mathfrak{S}\subset\mathfrak{N}$. Now if the ideal $\mathfrak{a}$ is semiprincipal, we have either $\mathfrak{a}=\mathfrak{a}(a)$ or $\mathfrak{a}=\mathfrak{a}'(a)$ for some element a. In the first case, we let b be an arbitrary element in A and write $b=ab+(b+ab)$. Since ab is in $\mathfrak{a}$, a proof that $b+ab$ is in $\mathfrak{a}'$ will lead to the result that b is in $\mathfrak{a}\vee\mathfrak{a}'$. If c is an arbitrary element in $\mathfrak{a}=\mathfrak{a}(a)$, we have $c(b+ab)=cb+(ca)b=cb+cb=0$ and thus conclude that $b+ab$ is in $\mathfrak{a}'$. Since b is arbitrary, we must have $\mathfrak{a}\vee\mathfrak{a}'=\mathfrak{e}$, so that $\mathfrak{a}$ is a simple ideal. In the second case we apply the result just obtained to write $\mathfrak{e}=\mathfrak{a}(a)\vee\mathfrak{a}'(a)\subset\mathfrak{a}''(a)\vee\mathfrak{a}'(a)=\mathfrak{a}'\vee\mathfrak{a}=\mathfrak{a}\vee\mathfrak{a}'$. It is then evident that $\mathfrak{a}\vee\mathfrak{a}'=\mathfrak{e}$ and that $\mathfrak{a}$ is simple. The proof of the relation $\mathfrak{P}^*\subset\mathfrak{S}$ is thus completed. If now $\mathfrak{a}$ is a simple ideal, we have $\mathfrak{a}''=\mathfrak{a}''\mathfrak{e}=\mathfrak{a}''(\mathfrak{a}\vee\mathfrak{a}')=\mathfrak{a}''\mathfrak{a}\vee\mathfrak{a}''\mathfrak{a}'=\mathfrak{a}\vee\mathfrak{o}=\mathfrak{a}$ in accordance with Theorems 18 and 21; and we conclude that $\mathfrak{a}$ is normal. The relation $\mathfrak{S}\subset\mathfrak{N}$ is thus established.

THEOREM 24. *The relation $\mathfrak{J}\neq\mathfrak{S}$ implies the relation $\mathfrak{J}\neq\mathfrak{N}$; in particular, if the ideal $\mathfrak{a}$ is not simple, the ideal $\mathfrak{a}\vee\mathfrak{a}'$ is not normal. In consequence, the relation $\mathfrak{J}=\mathfrak{N}$ implies the relation $\mathfrak{J}=\mathfrak{S}$.*

Since $(\mathfrak{a}\vee\mathfrak{a}')'=\mathfrak{a}'\mathfrak{a}''=\mathfrak{o}$ by Theorem 21, and since therefore $(\mathfrak{a}\vee\mathfrak{a}')''=\mathfrak{e}$, we see that $\mathfrak{a}\vee\mathfrak{a}'\neq\mathfrak{e}$ implies $\mathfrak{a}\vee\mathfrak{a}'\neq(\mathfrak{a}\vee\mathfrak{a}')''$.

THEOREM 25. *The relation* $\mathfrak{P}=\mathfrak{P}^*$ *implies the relation* $\mathfrak{P}=\mathfrak{S}$ *and hence also the relation* $\mathfrak{P}^*=\mathfrak{S}$. *In fact, the following assertions concerning a Boolean ring* A *are equivalent*:

(1) $\mathfrak{P}=\mathfrak{S}$; (2) $\mathfrak{P}=\mathfrak{P}^*$;

(3) *there exists an ideal* $\mathfrak{a}$ *such that* $\mathfrak{a}$ *and* $\mathfrak{a}'$ *are in* $\mathfrak{P}$;

(4) *the Boolean ring* A *has a unit* e.

It is evident that (1) implies (2), and also that (2) implies (3). We show that (3) implies (4). If $\mathfrak{a}=\mathfrak{a}(a)$ and $\mathfrak{a}'=\mathfrak{a}(b)$, then $\mathfrak{a}(a)\vee\mathfrak{a}(b)=\mathfrak{a}\vee\mathfrak{a}'=\mathfrak{e}$ in accordance with the relation $\mathfrak{P}\equiv\mathfrak{S}$. Hence an arbitrary element c is expressible in the form $c=a_1\vee b_1$ where $a_1<a$ and $b_1<b$, as we see by reference to Theorem 17. It follows that $c<a\vee b$. Since c is arbitrary, this relation identifies $a\vee b$ as the unit in A. We now complete the proof by showing that (4) implies (1). If $\mathfrak{a}$ is a simple ideal in a Boolean ring with unit e, the relation $\mathfrak{a}\vee\mathfrak{a}'=\mathfrak{e}$ shows that $e=a\vee b$ where a is in $\mathfrak{a}$ and b in $\mathfrak{b}$, in accordance with Theorem 17. If c is an arbitrary element in $\mathfrak{a}$, we obviously have $cb=0$ and hence $c=ce=c(a\vee b)=ca\vee cb=ca$, or, equivalently, $c<a$. Thus we see that $\mathfrak{a}\subset\mathfrak{a}(a)$. On the other hand it is evident that $\mathfrak{a}(a)\subset\mathfrak{a}$. We therefore conclude that $\mathfrak{a}=\mathfrak{a}(a)$, as we wished to do.

We may remark that in a Boolean ring without unit we have, in general, $\mathfrak{P}\neq\mathfrak{P}^*\neq\mathfrak{S}\neq\mathfrak{N}\neq\mathfrak{J}$; and that in a Boolean ring with unit we have, in general, $\mathfrak{P}=\mathfrak{P}^*=\mathfrak{S}\neq\mathfrak{N}\neq\mathfrak{J}$. When $\mathfrak{J}\neq\mathfrak{N}$ we may have either $\mathfrak{N}\neq\mathfrak{S}$ or $\mathfrak{N}=\mathfrak{S}$; but we have not been able to determine whether the relations $\mathfrak{P}\neq\mathfrak{P}^*$, $\mathfrak{P}^*=\mathfrak{S}$ are compatible or not. The various possibilities will be analyzed more fully on another occasion.

The conditions that an ideal be principal or semiprincipal, as given in Definition 8, are adequate for our purposes; but it is important for us to obtain conditions, other than those given in the definition, for an ideal to be simple or normal. The two theorems which follow present information on this topic.

THEOREM 26. *In order that an ideal* $\mathfrak{a}$ *in a Boolean ring* A *be simple, it is necessary and sufficient that the product* $\mathfrak{a}\mathfrak{a}(a)$ *be a principal ideal for every element* a *in* A.

We shall consider $\mathfrak{a}\mathfrak{a}(a)$ as an ideal in the subring $\mathfrak{a}(a)$ with unit a, recalling that the orthocomplement of $\mathfrak{a}\mathfrak{a}(a)$ relative to $\mathfrak{a}(a)$ is the ideal $\mathfrak{a}'\mathfrak{a}(a)$ in accordance with Theorem 22. If $\mathfrak{a}$ is simple relative to A, we have $\mathfrak{a}\mathfrak{a}(a)\vee\mathfrak{a}'\mathfrak{a}(a)=(\mathfrak{a}\vee\mathfrak{a}')\mathfrak{a}(a)=\mathfrak{e}\mathfrak{a}(a)=\mathfrak{a}(a)$ so that $\mathfrak{a}\mathfrak{a}(a)$ is simple relative to $\mathfrak{a}(a)$. By virtue of Theorem 25 we see that $\mathfrak{a}\mathfrak{a}(a)$ is principal relative to $\mathfrak{a}(a)$. Since $\mathfrak{a}\mathfrak{a}(a)$ is an ideal in A and since, considered as a Boolean ring, it has a unit by virtue of its character in $\mathfrak{a}(a)$, we conclude that there exists an element b such that

$\mathfrak{a}\mathfrak{a}(a) = \mathfrak{a}(b)$. Thus, when $\mathfrak{a}$ is simple, $\mathfrak{a}\mathfrak{a}(a)$ is principal. On the other hand, if $\mathfrak{a}\mathfrak{a}(a)$ is principal for every a in A, we write $\mathfrak{a}\mathfrak{a}(a) = \mathfrak{a}(b) \subset \mathfrak{a}(a)$. It follows that $\mathfrak{a}\mathfrak{a}(a)$ is principal, and hence simple, relative to $\mathfrak{a}(a)$. Thus we see that $\mathfrak{a}(a) = \mathfrak{a}\mathfrak{a}(a) \vee \mathfrak{a}'\mathfrak{a}(a) = (\mathfrak{a} \vee \mathfrak{a}')\mathfrak{a}(a) \subset \mathfrak{a} \vee \mathfrak{a}'$; in other words, that $\mathfrak{a} \vee \mathfrak{a}'$ contains a. Since a is arbitrary, we conclude that $\mathfrak{a} \vee \mathfrak{a}' = \mathfrak{e}$, or, in other words, that $\mathfrak{a}$ is simple, as we wished to prove.

THEOREM 27. *The following assertions concerning an ideal $\mathfrak{a}$ in a Boolean ring A are equivalent*:

(1) *$\mathfrak{a}$ is a normal ideal*;

(2) *$\mathfrak{a}$ is the orthocomplement of some ideal in A*;

(3) *$\mathfrak{a}$ is a product of semiprincipal ideals.*

In general, if $\mathfrak{a}$ is an arbitrary ideal, then $\mathfrak{a}''$ is the product of all the semiprincipal ideal divisors of $\mathfrak{a}$; and, in particular, a normal ideal is the product of all its semiprincipal ideal divisors. In the case of a Boolean ring with unit, the term "principal ideal" is to replace the term "semiprincipal ideal" in the preceding statements.

The equivalence of (1) and (2) is easily proved as follows: if $\mathfrak{a}$ is normal, then $\mathfrak{a} = \mathfrak{b}'$ for $\mathfrak{b} = \mathfrak{a}'$; and, if $\mathfrak{a} = \mathfrak{b}'$, then $\mathfrak{a} = \mathfrak{b}' = \mathfrak{b}''' = \mathfrak{a}''$. If $\mathfrak{a}$ is an arbitrary ideal, it has at least one semiprincipal ideal divisor, the ideal $\mathfrak{e}$. Hence the product of all the semiprincipal ideal divisors of $\mathfrak{a}$ exists and is an ideal $\mathfrak{b}$ which divides $\mathfrak{a}$. Now if $\mathfrak{c}$ is a semiprincipal ideal divisor of $\mathfrak{a}$, we have $\mathfrak{c}' \subset \mathfrak{a}'$, $\mathfrak{c} = \mathfrak{c}'' \supset \mathfrak{a}''$, so that $\mathfrak{c}$ is also a semiprincipal ideal divisor of $\mathfrak{a}''$. It follows that $\mathfrak{b} \supset \mathfrak{a}''$. On the other hand, if a is an arbitrary element in $\mathfrak{a}'$, we see that $\mathfrak{a}(a) \subset \mathfrak{a}'$, that $\mathfrak{a}'(a) \supset \mathfrak{a}'' \supset \mathfrak{a}$, and hence that $\mathfrak{a}'(a) \supset \mathfrak{b}$. The latter relation implies that $\mathfrak{a}(a)\mathfrak{b} = \mathfrak{o}$. Since a is arbitrary, we see that $\mathfrak{a}'\mathfrak{b} = \mathfrak{o}$ and hence that $\mathfrak{b} \subset \mathfrak{a}''$. It follows from this and the earlier inclusion relation that $\mathfrak{b} = \mathfrak{a}''$; in other words, that $\mathfrak{a}''$ is the product of all the semiprincipal ideal divisors of $\mathfrak{a}$. The special case where $\mathfrak{a}$ is normal, that is, where $\mathfrak{a} = \mathfrak{a}''$, is now obvious. If an ideal $\mathfrak{a}$ is the product of semiprincipal ideals, then we must have $\mathfrak{a} \supset \mathfrak{a}''$ since $\mathfrak{a}''$ is the product of *all* the semiprincipal ideal divisors of $\mathfrak{a}$; we now conclude by virtue of the relation $\mathfrak{a} \subset \mathfrak{a}''$ that $\mathfrak{a} = \mathfrak{a}''$ and hence that $\mathfrak{a}$ is normal. The preceding results show the equivalence of (1) and (3). The final statement of the theorem follows immediately from Theorem 25.

We shall turn now to an examination of the behavior of the various classes of ideals under the operations of addition and multiplication. As a preliminary to our first theorem in this connection, we must make a few remarks concerning congruences in arbitrary algebraic systems. In any algebraic system, the fundamental relation of equality has to be taken as one of the undefined concepts and must be connected with the operations of the

system by suitable postulates. It is sufficient to assume, first, that the relation of equality is a reflexive, symmetric, and transitive dyadic relation, and, second, that in terms of this relation the following law of substitution holds: for each of the operations of the system, the substitution of equal operands for given operands respectively replaces the element resulting from the application of that operation by an equal element. Now any dyadic relation which has these same properties and which holds whenever the fundamental relation of equality holds may be called a congruence in the given algebraic system. The fundamental equality is itself a congruence. By simple inductive arguments it can be shown that, for any given relation of congruence, the following general rule of substitution is valid: if an element a is obtained as a polynomial in terms of elements of the system, then the substitution of respectively congruent elements in this polynomial yields an element congruent to a. Evidently, any relation of congruence in an algebraic system may be used to replace the fundamental relation of equality; if this be done, the system is converted into a new system which is easily seen to be homomorphic to the old. Conversely, any system homomorphic to the given one is isomorphic to a system obtained in this way by the use of an appropriate congruence: when the homomorphism is given, the associated congruence is obtained by defining two elements to be congruent if and only if they are carried by the homomorphism into equal elements of the homomorphic system. The following theorem is stated with these general remarks as a background.

THEOREM 28. *The dyadic relation C defined between the elements of the class $\mathfrak{J}$ of all ideals in a Boolean ring A by setting $\mathfrak{a} \equiv \mathfrak{b}$ if $\mathfrak{a}' = \mathfrak{b}'$, is a congruence in the algebraic system consisting of the class $\mathfrak{J}$ and the operations of unrestricted addition and finite multiplication. Each class of mutually congruent elements in $\mathfrak{J}$ contains one and only one normal ideal as an element, in the following sense: if $\mathfrak{a}$ is any ideal, then $\mathfrak{a}''$ is a normal ideal such that $\mathfrak{a} \equiv \mathfrak{a}''$; and, if $\mathfrak{a}$ and $\mathfrak{b}$ are normal ideals such that $\mathfrak{a} \equiv \mathfrak{b}$, then $\mathfrak{a} = \mathfrak{b}$. The algebraic system $\mathfrak{J}^C$ consisting of the class $\mathfrak{J}$ with the congruence C as the fundamental relation of equality and the operations of finite addition and finite multiplication is a Boolean algebra with unit in accordance with Theorem 3.*†

Most of the properties of the relation C asserted in the theorem are easily verified. Thus $\mathfrak{a} = \mathfrak{b}$ implies $\mathfrak{a}' = \mathfrak{b}'$ and hence $\mathfrak{a} \equiv \mathfrak{b}$; in particular, $\mathfrak{a} \equiv \mathfrak{a}$. We see that $\mathfrak{a} \equiv \mathfrak{b}$ implies $\mathfrak{b} \equiv \mathfrak{a}$, since $\mathfrak{a}' = \mathfrak{b}'$ implies $\mathfrak{b}' = \mathfrak{a}'$; and that $\mathfrak{a} \equiv \mathfrak{b}$ and $\mathfrak{b} \equiv \mathfrak{c}$

† This theorem bears a close formal or structural relation to a general proposition about the logic of Brouwer in the symbolic statement of Heyting, Sitzungsberichte der Preussischen Akademie der Wissenschaften, 1930, pp. 42–56. The proposition in question was given by Glivenko, Académie Royale de Belgique, Bulletins des Sciences, (5), vol. 15 (1929), pp. 183–188.

imply $\mathfrak{a} \equiv \mathfrak{c}$, since $\mathfrak{a}' = \mathfrak{b}'$ and $\mathfrak{b}' = \mathfrak{c}'$ imply $\mathfrak{a}' = \mathfrak{c}'$. If $\mathfrak{A}$ and $\mathfrak{B}$ are subclasses of $\mathfrak{J}$ in one-to-one correspondence in such a manner that corresponding elements $\mathfrak{a}$ and $\mathfrak{b}$ are in the relation $\mathfrak{a} \equiv \mathfrak{b}$, we see that

$$(\underset{\mathfrak{a}\in\mathfrak{A}}{\mathrm{S}}\,\mathfrak{a})' = \underset{\mathfrak{a}\in\mathfrak{A}}{\mathrm{P}}\,\mathfrak{a}' = \underset{\mathfrak{b}\in\mathfrak{B}}{\mathrm{P}}\,\mathfrak{b}' = (\underset{\mathfrak{b}\in\mathfrak{B}}{\mathrm{S}}\,\mathfrak{b})'$$

in accordance with Theorem 21, and hence that

$$\underset{\mathfrak{a}\in\mathfrak{A}}{\mathrm{S}}\,\mathfrak{a} \equiv \underset{\mathfrak{b}\in\mathfrak{B}}{\mathrm{S}}\,\mathfrak{b}.$$

In particular, we note that $\mathfrak{a} \equiv \mathfrak{c}$ and $\mathfrak{b} \equiv \mathfrak{d}$ imply $\mathfrak{a} \vee \mathfrak{b} \equiv \mathfrak{c} \vee \mathfrak{d}$.

The proof that $\mathfrak{a} \equiv \mathfrak{c}$ and $\mathfrak{b} \equiv \mathfrak{d}$ imply $\mathfrak{ab} \equiv \mathfrak{cd}$ is more difficult. We shall first prove that $\mathfrak{a} \equiv \mathfrak{e}$ and $\mathfrak{b} \equiv \mathfrak{e}$ imply $\mathfrak{ab} \equiv \mathfrak{e}$, or, equivalently, that $\mathfrak{a}' = \mathfrak{o}$ and $\mathfrak{b}' = \mathfrak{o}$ imply $(\mathfrak{ab})' = \mathfrak{o}$. Let a be an arbitrary element in $(\mathfrak{ab})'$, b an arbitrary element in $\mathfrak{a}(a)\mathfrak{a}$, and c an arbitrary element in $\mathfrak{a}(b)\mathfrak{b}$. It is then clear that $\mathfrak{a}(c) \subset \mathfrak{a}(b)\mathfrak{b} \subset \mathfrak{a}(a)\mathfrak{ab} \subset (\mathfrak{ab})'\mathfrak{ab} = \mathfrak{o}$ and hence that $c = 0$. Since c was an arbitrary element in $\mathfrak{a}(b)\mathfrak{b}$, we conclude that $\mathfrak{a}(b)\mathfrak{b} = \mathfrak{o}$ and hence that $\mathfrak{a}(b) \subset \mathfrak{b}' = \mathfrak{o}$, $b = 0$. Since b was an arbitrary element in $\mathfrak{a}(a)\mathfrak{a}$, we conclude that $\mathfrak{a}(a)\mathfrak{a} = \mathfrak{o}$ and hence that $\mathfrak{a}(a) \subset \mathfrak{a}' = \mathfrak{o}$, $a = 0$. Since a was an arbitrary element in $(\mathfrak{ab})'$, we conclude that $(\mathfrak{ab})' = \mathfrak{o}$, as we wished to prove. We next prove that $\mathfrak{a}''\mathfrak{b}''(\mathfrak{ab})' = \mathfrak{o}$. We begin by writing $\mathfrak{a}''\mathfrak{b}''(\mathfrak{ab})' = (\mathfrak{a}' \vee \mathfrak{b}')'(\mathfrak{ab})' = (\mathfrak{a}' \vee \mathfrak{b}' \vee \mathfrak{ab})'$ in accordance with Theorem 21. We then observe that $\mathfrak{a}' \vee \mathfrak{b}' \vee \mathfrak{ab} \supset \mathfrak{a}'(\mathfrak{b} \vee \mathfrak{b}') \vee \mathfrak{ab}' \vee \mathfrak{ab} = (\mathfrak{a} \vee \mathfrak{a}')(\mathfrak{b} \vee \mathfrak{b}')$ and hence that $(\mathfrak{a}' \vee \mathfrak{b}' \vee \mathfrak{ab})' \subset [(\mathfrak{a} \vee \mathfrak{a}')(\mathfrak{b} \vee \mathfrak{b}')]'$. In view of the relations $(\mathfrak{a} \vee \mathfrak{a}')' = \mathfrak{a}'\mathfrak{a}'' = \mathfrak{o}$, $(\mathfrak{b} \vee \mathfrak{b}')' = \mathfrak{b}'\mathfrak{b}'' = \mathfrak{o}$, we can conclude that $\mathfrak{a}''\mathfrak{b}''(\mathfrak{ab})' = \mathfrak{o}$ by virtue of the preceding results. We now observe that $(\mathfrak{a}''\mathfrak{b}'')' \supset (\mathfrak{ab})'$ as a consequence of the relation just established; and that $(\mathfrak{a}''\mathfrak{b}'')' \subset (\mathfrak{ab})'$ in consequence of the relation $\mathfrak{a}''\mathfrak{b}'' \supset \mathfrak{ab}$. We thus see that $(\mathfrak{a}''\mathfrak{b}'')' = (\mathfrak{ab})'$, or, equivalently, that $\mathfrak{a}''\mathfrak{b}'' \equiv \mathfrak{ab}$. The final step of the proof is simple. If $\mathfrak{a} \equiv \mathfrak{c}$ and $\mathfrak{b} \equiv \mathfrak{d}$, then we have $\mathfrak{a}' = \mathfrak{c}'$, $\mathfrak{a}'' = \mathfrak{c}''$ and $\mathfrak{b}' = \mathfrak{b}'$, $\mathfrak{b}'' = \mathfrak{b}''$; hence we have $\mathfrak{a}''\mathfrak{b}'' = \mathfrak{c}''\mathfrak{b}''$ and $\mathfrak{a}''\mathfrak{b}'' \equiv \mathfrak{c}''\mathfrak{b}''$; and finally, from the relations $\mathfrak{a}''\mathfrak{b}'' \equiv \mathfrak{ab}$ and $\mathfrak{c}''\mathfrak{b}'' \equiv \mathfrak{cb}$ we conclude that $\mathfrak{ab} \equiv \mathfrak{cb}$. We can show by examples that this result cannot be extended to the case of unrestricted multiplication.

The assertion concerning the classes of mutually congruent elements is obvious.

Having shown that the relation C is a congruence in the indicated sense, we may use it to replace the fundamental equality in $\mathfrak{J}$ in accordance with the general remarks above. When we restrict consideration to the finite operations $\vee$ and $\cdot$ alone, we obtain the system $\mathfrak{J}^c$. We wish to identify this system as a Boolean algebra with $\mathfrak{e}$ as unit, by showing that it has the properties (1_1), (3_1), (3_2), (4_1), (5), (6_1) and (6_2) of Theorem 3. By combining the properties of the congruence C with the results given in Theorem 18, we establish

all the desired properties, except (5), without any difficulty. To prove (5), we show that the relations $\mathfrak{x} \vee \mathfrak{a} \equiv \mathfrak{e}$, $\mathfrak{x}\mathfrak{a} \equiv \mathfrak{o}$ have a solution $\mathfrak{x} \equiv \mathfrak{a}'$. To this end we have merely to note that $(\mathfrak{a}' \vee \mathfrak{a})' = \mathfrak{a}''\mathfrak{a}' = \mathfrak{o} = \mathfrak{e}'$, $\mathfrak{a}'\mathfrak{a} = \mathfrak{o}$, and hence that $\mathfrak{a}' \vee \mathfrak{a} \equiv \mathfrak{e}$, $\mathfrak{a}'\mathfrak{a} \equiv \mathfrak{o}$. To convert $\mathfrak{J}^c$ into a Boolean ring we must replace the operation $\vee$ by an operation $+$ such that $\mathfrak{a}+\mathfrak{b} \equiv \mathfrak{a}\mathfrak{b}' \vee \mathfrak{a}'\mathfrak{b}$, in accordance with Theorem 2. Evidently, it is most convenient to take $\mathfrak{a}+\mathfrak{b}$ as the ideal $\mathfrak{a}\mathfrak{b}' \vee \mathfrak{a}'\mathfrak{b}$, rather than as some other ideal congruent to the latter.

In discussing the class of all normal ideals, it will be necessary to replace addition by a new operation. We therefore give the following formal definition:

DEFINITION 9. *If $\mathfrak{B}$ is a non-void class of ideals $\mathfrak{a}$, the ideal $(S_{\mathfrak{a}\in\mathfrak{B}}\mathfrak{a})''$ is called the normalized sum of the ideals $\mathfrak{a}$ in $\mathfrak{B}$ and is denoted by $S''_{\mathfrak{a}\in\mathfrak{B}}\mathfrak{a}$; the normalized sum of ideals $\mathfrak{a}$ and $\mathfrak{b}$ is denoted by $\mathfrak{a} \vee \mathfrak{b}$. The operation of forming the normalized sum is called normalized addition.*

In terms of this definition, we have

THEOREM 29. *The normalized sum and the product of normal ideals are normal ideals; but a finite sum of normal ideals is not necessarily normal. The normalized sum of arbitrary ideals is the least normal ideal containing all the summands. Within the class $\mathfrak{N}$ of all normal ideals the operations of normalized addition and multiplication have the following properties:*

(1) *if $\mathfrak{B}$ is a non-void class of non-void subclasses $\mathfrak{B}$ of $\mathfrak{N}$, then*

$$S''_{\mathfrak{B}\in\mathfrak{B}} \; (S''\mathfrak{a}) = S''\mathfrak{a},$$

where $\mathfrak{C}$ is the union $\Sigma_{\mathfrak{B}\in\mathfrak{B}}\mathfrak{B}$;

(2) *if $\mathfrak{B}$, $\mathfrak{B}$, and $\mathfrak{C}$ have the same significance as in (1), then*

$$P_{\mathfrak{B}\in\mathfrak{B}} \; (P\,\mathfrak{a}) = P\,\mathfrak{a};$$

(3) *if $\mathfrak{b}$ is any normal ideal and $\mathfrak{B}$ is any non-void subclass of $\mathfrak{N}$, then*

$$\mathfrak{b}(S''\mathfrak{a}) = S''(\mathfrak{b}\mathfrak{a});$$

(4) *if $\mathfrak{b}$ and $\mathfrak{B}$ have the same significance as in (3), then*

$$P_{\mathfrak{a}\in\mathfrak{B}}(\mathfrak{b} \vee \mathfrak{a}) = \mathfrak{b} \vee P\,\mathfrak{a};$$

(5) *if $\mathfrak{B}$ is any non-void subclass of $\mathfrak{N}$, then*

$$(S''\mathfrak{a})' = P\,\mathfrak{a}';$$

(6) *if* $\mathfrak{B}$ *has the same significance as in* (5), *then*

$$(\underset{\mathfrak{a}\in\mathfrak{B}}{\mathrm{P}}\,\mathfrak{a})' = \underset{\mathfrak{a}\in\mathfrak{B}}{\mathrm{S}}''\mathfrak{a}'.$$

Under the finite operations ∇ *and* $\cdot$ *alone, the system* $\mathfrak{N}$ *is a Boolean algebra isomorphic to the system* $\mathfrak{Z}^c$ *of Theorem* 28 *by virtue of the correspondence* $\mathfrak{a}\longleftrightarrow\mathfrak{a}''$. *This algebra has the property that its normal ideals are all principal.*

In view of Theorem 27(2) and Definition 9, it is evident that the normalized sum of ideals is always a normal ideal. In view of Theorem 27(3), it is likewise evident that the product of normal ideals is a normal ideal: since each factor is a product of semiprincipal ideals, the product is also a product of semiprincipal ideals by Theorem 15(11) and Theorem 18. On the other hand, if $\mathfrak{a}$ is a normal ideal which is not simple, the sum $\mathfrak{a}\vee\mathfrak{a}'$ is not normal, as we showed in Theorem 24, in spite of the fact that $\mathfrak{a}$ and $\mathfrak{a}'$ are both normal.

If $\mathfrak{a}$ is an arbitrary ideal, then $\mathfrak{a}''$ is a normal ideal containing $\mathfrak{a}$. If $\mathfrak{b}$ is a normal ideal containing $\mathfrak{a}$, then $\mathfrak{b}$ is the product of semiprincipal ideal divisors of $\mathfrak{a}$ and thus contains $\mathfrak{a}''$, in accordance with Theorem 27. Thus $\mathfrak{a}''$ is the least normal ideal containing $\mathfrak{a}$. By comparing this result with Definition 9, we see that the normalized sum of ideals is the least normal ideal containing all the summands.

Some of the properties (1)–(6) have already been established. Thus (2) has been proved in Theorem 15 (11) and Theorem 18, and (5) follows from Theorem 21(4) and Definition 9 by the use of the relation $\mathfrak{a}'''=\mathfrak{a}'$, since we have

$$(\underset{\mathfrak{a}\in\mathfrak{B}}{\mathrm{S}}''\mathfrak{a})' = (\underset{\mathfrak{a}\in\mathfrak{B}}{\mathrm{S}}\,\mathfrak{a})''' = (\underset{\mathfrak{a}\in\mathfrak{B}}{\mathrm{S}}\,\mathfrak{a})' = \underset{\mathfrak{a}\in\mathfrak{B}}{\mathrm{P}}\,\mathfrak{a}'.$$

We can now deduce (6) from (5) by virtue of the relations

$$(\underset{\mathfrak{a}\in\mathfrak{B}}{\mathrm{P}}\,\mathfrak{a})' = (\underset{\mathfrak{a}\in\mathfrak{B}}{\mathrm{P}}\,\mathfrak{a}'')' = (\underset{\mathfrak{a}\in\mathfrak{B}}{\mathrm{S}}''\mathfrak{a}')'' = \underset{\mathfrak{a}\in\mathfrak{B}}{\mathrm{S}}''\mathfrak{a}',$$

since the ideals in $\mathfrak{B}$ and all normalized sums are normal ideals. We can establish (1) from the corresponding property of ordinary sums already established in Theorem 15(10) and Theorem 18 by using the results of Theorem 28 concerning the congruence C: from Definition 9 it is evident that any normalized sum is congruent to the corresponding ordinary sum; in consequence, the two members of (1) above are congruent, the corresponding ordinary sums being equal; and finally, since the members are both normal ideals, their congruence implies their equality. In a similar way, we establish (3) from the corresponding relation for ordinary sums, already proved in Theorem 18(12): the two members of (3) are congruent and, being normal ideals,

are therefore equal. With the help of (5) and (6) it is now easy to deduce (4) from (3), as follows:

$$\mathop{P}_{\mathfrak{a}\in\mathfrak{B}}(\mathfrak{b}\nabla\mathfrak{a}) = \mathop{P}_{\mathfrak{a}\in\mathfrak{B}}(\mathfrak{b}''\nabla\mathfrak{a}'') = \mathop{P}_{\mathfrak{a}\in\mathfrak{B}}(\mathfrak{b}'\mathfrak{a}') = (\mathop{S}_{\mathfrak{a}\in\mathfrak{B}}''(\mathfrak{b}'\mathfrak{a}'))' = (\mathfrak{b}'(\mathop{S}_{\mathfrak{a}\in\mathfrak{B}}''\mathfrak{a}'))'$$

$$= (\mathfrak{b}'(\mathop{P}_{\mathfrak{a}\in\mathfrak{B}}\mathfrak{a})')' = \mathfrak{b}''\nabla(\mathop{P}_{\mathfrak{a}\in\mathfrak{B}}\mathfrak{a})'' = \mathfrak{b}\nabla\mathop{P}_{\mathfrak{a}\in\mathfrak{B}}\mathfrak{a},$$

the ideals $\mathfrak{a}$, $\mathfrak{b}$, and $P_{\mathfrak{a}\in\mathfrak{B}}\mathfrak{a}$ being normal.

In order to show that under the finite operations ∇ and $\cdot$ the system $\mathfrak{R}$ is a Boolean algebra, we must verify properties (1_1), (3_1), (3_2), (4_1), (5), (6_1), (6_2) of Theorem 3. Now (1_1), (4_1), (6_1), and (6_2) are evident from the known properties of the operation $\vee$ and the relation $\mathfrak{a}\nabla\mathfrak{b} = (\mathfrak{a}\vee\mathfrak{b})''$. Properties (3_1) and (3_2) follow at once from (3) above and the commutative law for multiplication. To establish (5) we show that the equations $\mathfrak{x}\nabla\mathfrak{a} = \mathfrak{e}$, $\mathfrak{x}\mathfrak{a} = \mathfrak{o}$, where $\mathfrak{o}$ and $\mathfrak{e}$ are known to be normal ideals and $\mathfrak{a}$ is assumed to be a normal ideal, have as a solution the normal ideal $\mathfrak{x} = \mathfrak{a}'$: we have only to note the relations $\mathfrak{a}'\nabla\mathfrak{a} = (\mathfrak{a}'\vee\mathfrak{a})'' = (\mathfrak{a}''\mathfrak{a}')' = \mathfrak{o}' = \mathfrak{e}$, $\mathfrak{a}'\mathfrak{a} = \mathfrak{o}$. In order to convert this Boolean algebra into a Boolean ring, we have only to introduce the operation $+$ defined by the equation $\mathfrak{a}+\mathfrak{b} = \mathfrak{a}\mathfrak{b}'\nabla\mathfrak{a}'\mathfrak{b}$, in accordance with Theorem 2. It is now easily verified that the correspondence $\mathfrak{a}\longleftrightarrow\mathfrak{a}''$ sets up an isomorphism between the Boolean algebras $\mathfrak{Z}^c$ and $\mathfrak{R}$. This correspondence is biunivocal on account of Theorem 28, which shows that $\mathfrak{a}''$ is a normal ideal congruent to $\mathfrak{a}$ and that $\mathfrak{a}\equiv\mathfrak{b}$ if and only if $\mathfrak{a}'' = \mathfrak{b}''$. The correspondent of $\mathfrak{a}\vee\mathfrak{b}$ is given by $(\mathfrak{a}\vee\mathfrak{b})'' = (\mathfrak{a}'\mathfrak{b}')' = \mathfrak{a}''\nabla\mathfrak{b}''$ in accordance with Theorem 21(4) and (6) above. The correspondent of $\mathfrak{a}\mathfrak{b}$ is found to be $(\mathfrak{a}\mathfrak{b})'' = \mathfrak{a}''\mathfrak{b}''$, by the following reasoning: the normal ideal $(\mathfrak{a}\mathfrak{b})''$ is congruent to $\mathfrak{a}\mathfrak{b}$ and hence also to $\mathfrak{a}''\mathfrak{b}''$; since $\mathfrak{a}''\mathfrak{b}''$ is normal, we must have $(\mathfrak{a}\mathfrak{b})'' = \mathfrak{a}''\mathfrak{b}''$ in accordance with Theorem 28. These results establish the indicated isomorphism.

It remains for us to prove that the Boolean algebra $\mathfrak{R}$ has the special property that its normal ideals are all principal. Since $\mathfrak{R}$ has the ideal $\mathfrak{e}$ as its unit, all simple ideals of the algebra $\mathfrak{R}$ are principal; in other words, a simple ideal of $\mathfrak{R}$ is characterized by an element $\mathfrak{a}$ in $\mathfrak{R}$ such that $\mathfrak{b}$ in $\mathfrak{R}$ belongs to the ideal if and only if $\mathfrak{b}\subset\mathfrak{a}$. This result follows from Theorem 25 and applies in particular to semiprincipal ideals of $\mathfrak{R}$. Now a normal ideal of $\mathfrak{R}$ is by virtue of Theorem 27 the product of semiprincipal ideals and hence, by virtue of the result just noted, the product of principal ideals. If $\mathfrak{B}$ is the class of elements $\mathfrak{a}$ generating the various factors in a product of principal ideals, the element $P_{\mathfrak{a}\in\mathfrak{B}}\mathfrak{a}$ belongs to $\mathfrak{R}$ and is contained in every element $\mathfrak{a}$ in $\mathfrak{B}$; and any element of $\mathfrak{R}$ belonging to the product is contained in every element $\mathfrak{a}$, and hence also in the element $P_{\mathfrak{a}\in\mathfrak{B}}\mathfrak{a}$. Thus the product of principal ideals in the

algebra $\mathfrak{N}$ is itself the principal ideal generated by the indicated element $P_{\mathfrak{a}\in\mathfrak{B}}\mathfrak{a}$. It follows that every normal ideal of $\mathfrak{N}$ is principal.

Theorem 29 can be summarized briefly by pointing out that under the operations S'', P, and ' the class $\mathfrak{N}$ has *some* of the chief formal properties of an algebra of classes with the corresponding operations of forming *unrestricted* unions, products, and complements respectively. That it does not have *all* the formal properties of such an algebra we shall see later in the present paper.

THEOREM 30. *The class $\mathfrak{S}$ of all simple ideals in a Boolean ring A is a Boolean subring, with $\mathfrak{e}$ as its unit, of the Boolean algebras $\mathfrak{J}^c$ and $\mathfrak{N}$ of Theorems 28 and 29 respectively. The application of the operations of finite addition, finite normalized addition, finite multiplication, and orthocomplementation to simple ideals yields simple ideals; in particular, if $\mathfrak{a}$ and $\mathfrak{b}$ are simple ideals, $\mathfrak{a}\dot{\vee}\mathfrak{b}=\mathfrak{a}\vee\mathfrak{b}$.*

If $\mathfrak{a}$ and $\mathfrak{b}$ are simple ideals, then $\mathfrak{a}\vee\mathfrak{b}$ is a simple ideal: for $(\mathfrak{a}\vee\mathfrak{b})\vee(\mathfrak{a}\vee\mathfrak{b})'=\mathfrak{a}\vee\mathfrak{b}\vee\mathfrak{a}'\mathfrak{b}'\supset\mathfrak{a}(\mathfrak{b}\vee\mathfrak{b}')\vee\mathfrak{a}'\mathfrak{b}\vee\mathfrak{a}'\mathfrak{b}'=(\mathfrak{a}\vee\mathfrak{a}')(\mathfrak{b}\vee\mathfrak{b}')=\mathfrak{e}\mathfrak{e}=\mathfrak{e}$ and hence $(\mathfrak{a}\vee\mathfrak{b})\vee(\mathfrak{a}\vee\mathfrak{b})'=\mathfrak{e}$. It follows also that $\mathfrak{a}\vee\mathfrak{b}$ is normal and that $\mathfrak{a}\vee\mathfrak{b}=(\mathfrak{a}\vee\mathfrak{b})''=\mathfrak{a}\dot{\vee}\mathfrak{b}$. If $\mathfrak{a}$ is a simple ideal, then $\mathfrak{a}'$ is a simple ideal: for $\mathfrak{a}'\vee\mathfrak{a}''\supset\mathfrak{a}'\vee\mathfrak{a}=\mathfrak{e}$ and hence $\mathfrak{a}'\vee\mathfrak{a}''=\mathfrak{e}$. The results just proved show that $\mathfrak{a}\mathfrak{b}$ and $\mathfrak{a}+\mathfrak{b}=\mathfrak{a}\mathfrak{b}'\dot{\vee}\mathfrak{a}'\mathfrak{b}=\mathfrak{a}\mathfrak{b}'\vee\mathfrak{a}'\mathfrak{b}$ are simple ideals whenever $\mathfrak{a}$ and $\mathfrak{b}$ are simple ideals: for $\mathfrak{a}\mathfrak{b}=\mathfrak{a}''\mathfrak{b}''=(\mathfrak{a}'\vee\mathfrak{b}')'$; and $\mathfrak{a}\mathfrak{b}'\dot{\vee}\mathfrak{a}'\mathfrak{b}=\mathfrak{a}\mathfrak{b}'\vee\mathfrak{a}'\mathfrak{b}$ whenever $\mathfrak{a}\mathfrak{b}'$ and $\mathfrak{a}'\mathfrak{b}$ are simple. Hence $\mathfrak{S}$ is a subring of $\mathfrak{J}^c$ and also of $\mathfrak{N}$. The ideal $\mathfrak{e}$ is in $\mathfrak{S}$ and is obviously its unit.

THEOREM 31. *The class $\mathfrak{P}$ of all principal ideals in a Boolean ring A is a Boolean subring of $\mathfrak{N}$ and an ideal in $\mathfrak{S}$; it is isomorphic to the Boolean ring A in accordance with the following relations.*

(1) $\mathfrak{a}(a)=\mathfrak{a}(b)$ *if and only if* $a=b$;

(2) $\mathfrak{a}(a+b)=\mathfrak{a}(a)+\mathfrak{a}(b)=\mathfrak{a}(a)\,\mathfrak{a}'(b)\vee\mathfrak{a}'(a)\mathfrak{a}(b)$;

(3) $\mathfrak{a}(a\vee b)=\mathfrak{a}(a)\vee\mathfrak{a}(b)$; (4) $\mathfrak{a}(ab)=\mathfrak{a}(a)\mathfrak{a}(b)$.

If the Boolean ring A has a unit e, then $\mathfrak{S}=\mathfrak{P}$ and $\mathfrak{a}(a')=\mathfrak{a}'(a)$.

The class $\mathfrak{P}$ is evidently non-void since it contains $\mathfrak{o}=\mathfrak{a}(0)$. In showing that it is an ideal in the Boolean algebra $\mathfrak{S}$, we shall establish properties (3) and (4) above. Theorem 14 shows that the element $a\vee b$ is contained in the ideal $\mathfrak{a}(a)\vee\mathfrak{a}(b)$ and hence that $\mathfrak{a}(a\vee b)\subset\mathfrak{a}(a)\vee\mathfrak{a}(b)$. On the other hand the relations $a<a\vee b$ and $b<a\vee b$ imply that $\mathfrak{a}(a)\subset\mathfrak{a}(a\vee b)$, $\mathfrak{a}(b)\subset\mathfrak{a}(a\vee b)$ and hence that $\mathfrak{a}(a)\vee\mathfrak{a}(b)\subset\mathfrak{a}(a\vee b)$. It follows that (3) is valid. Theorem 26 shows immediately that the ideal $\mathfrak{a}(a)\mathfrak{a}$ is principal, whatever the simple ideal $\mathfrak{a}$. In particular, $\mathfrak{a}(a)\mathfrak{a}(b)$ is principal. Theorem 16 thus shows that $\mathfrak{P}$ is an ideal in $\mathfrak{S}$ and hence a subring in $\mathfrak{N}$. Theorem 18 shows that ab is in the ideal $\mathfrak{a}(a)\mathfrak{a}(b)$ and hence that $\mathfrak{a}(ab)\subset\mathfrak{a}(a)\mathfrak{a}(b)$. On the other hand, the relations $c<a$

and $d<b$ imply $cd<ab$ and hence $\mathfrak{a}(a)\mathfrak{a}(b) \subset \mathfrak{a}(ab)$. It follows that (4) is valid. To prove (1), we note that $a=b$ obviously implies $\mathfrak{a}(a)=\mathfrak{a}(b)$; and that $\mathfrak{a}(a)=\mathfrak{a}(b)$ implies $a<b$, $b<a$ and hence $a=b$. The relations (1), (3), and (4) imply that the correspondence $a \longleftrightarrow \mathfrak{a}(a)$ sets up an isomorphism between A and $\mathfrak{P}$ with respect to the operations $\vee$ and $\cdot$ given in these systems. In order to extend this isomorphism to the operation $+$, we use (3) and (4) to prove (2). When $ab=0$ we have $a+b=a \vee b$ and $\mathfrak{a}(a)\mathfrak{a}(b)=\mathfrak{a}(ab)=\mathfrak{a}(0)=\mathfrak{o}$; and we therefore see that $\mathfrak{a}(a+b)=\mathfrak{a}(a \vee b)=\mathfrak{a}(a) \vee \mathfrak{a}(b)=\mathfrak{a}(a)+\mathfrak{a}(b)$. When a and b are arbitrary we apply this result to the elements $a+b$ and ab, which have the property $(a+b)ab=0$, to write $\mathfrak{a}(a+b)+\mathfrak{a}(ab)=\mathfrak{a}(a+b+ab)=\mathfrak{a}(a \vee b)$ $=\mathfrak{a}(a) \vee \mathfrak{a}(b)=\mathfrak{a}(a)+\mathfrak{a}(b)+\mathfrak{a}(a)\mathfrak{a}(b)=\mathfrak{a}(a)+\mathfrak{a}(b)+\mathfrak{a}(ab)$; and we conclude that $\mathfrak{a}(a+b)=\mathfrak{a}(a)+\mathfrak{a}(b)$, as we wished to show. In case A has a unit e, we know that $\mathfrak{P}$ coincides with $\mathfrak{S}$. By simple calculations, we find that $\mathfrak{a}(a')=\mathfrak{a}(a+e)=\mathfrak{a}(a)+\mathfrak{a}(e)=\mathfrak{a}(a)+\mathfrak{e}=\mathfrak{a}(a)\mathfrak{e}' \vee \mathfrak{a}'(a)\mathfrak{e}=\mathfrak{a}'(a)$, as stated in the theorem.

THEOREM 32. *The class $\mathfrak{P}^*$ of all semiprincipal ideals in a Boolean ring A is a subring of $\mathfrak{S}$, with $\mathfrak{e}$ as its unit, isomorphic to the Boolean ring B of Theorem 1; $\mathfrak{P}$ and A are ideals in $\mathfrak{P}^*$ and B respectively. The operations $\vee$, $\cdot$, $+$, and $'$ in the system $\mathfrak{S}$ apply to elements of $\mathfrak{P}^*$ in the manner indicated by the following rules:*

$$(1_1)\quad \mathfrak{a}(a) \vee \mathfrak{a}(b) = \mathfrak{a}(a \vee b); \qquad (1_2)\quad \mathfrak{a}(a) \vee \mathfrak{a}'(b) = \mathfrak{a}'(b + ab);$$

$$(1_3)\quad \mathfrak{a}'(a) \vee \mathfrak{a}'(b) = \mathfrak{a}'(ab);$$

$$(2_1)\quad \mathfrak{a}(a)\mathfrak{a}(b) = \mathfrak{a}(ab); \qquad (2_2)\quad \mathfrak{a}(a)\mathfrak{a}'(b) = \mathfrak{a}(a + ab);$$

$$(2_3)\quad \mathfrak{a}'(a)\mathfrak{a}'(b) = \mathfrak{a}'(a \vee b);$$

$$(3_1)\quad \mathfrak{a}(a) + \mathfrak{a}(b) = \mathfrak{a}(a + b); \qquad (3_2)\quad \mathfrak{a}(a) + \mathfrak{a}'(b) = \mathfrak{a}'(a + b);$$

$$(3_3)\quad \mathfrak{a}'(a) + \mathfrak{a}'(b) = \mathfrak{a}(a + b);$$

$$(4_1)\quad \mathfrak{a}'(a) \text{ is in } \mathfrak{P}^*; \qquad (4_2)\quad (\mathfrak{a}'(a))' = \mathfrak{a}(a).$$

In the case where $\mathfrak{P} \neq \mathfrak{P}^*$, $\mathfrak{P}$ is not a normal ideal in $\mathfrak{P}^*$.

Before giving proofs of the various numbered relations, we shall discuss their consequences. The class $\mathfrak{P}^*$ is non-void since it contains the ideals $\mathfrak{o}$ and $\mathfrak{e}$; it is a subring of $\mathfrak{S}$ by virtue of relations (2_1), (2_2), (2_3), (3_1), (3_2), and (3_3); and it obviously has $\mathfrak{e}$ as its unit. Since $\mathfrak{P}$ is an ideal in $\mathfrak{S}$ by Theorem 31, it is also an ideal in the subring $\mathfrak{P}^*$. When A has a unit, $\mathfrak{P}=\mathfrak{P}^*$ in accordance with Theorem 25, and the Boolean ring B of Theorem 1 coincides with A; the isomorphism between $\mathfrak{P}^*$ and B is thus a consequence of the isomorphism between $\mathfrak{P}$ and A established in Theorem 31. When A has no unit, $\mathfrak{P} \neq \mathfrak{P}^*$

and $A \neq B$. In this case we set up an isomorphism between $\mathfrak{P}^*$ and the Boolean ring of pairs (a, α) described in the proof of Theorem 1, and hence between $\mathfrak{P}^*$ and B. The necessary correspondence is indicated in the relations $\mathfrak{a}(a) \longleftrightarrow (a, 0)$, $\mathfrak{a}'(a) \longleftrightarrow (a, \epsilon)$. With the help of the relations (2_1), (2_2), (2_3), (3_1), (3_2), and (3_3), it is easily seen that this correspondence yields the desired isomorphism. It should be observed that the isomorphism between $\mathfrak{P}^*$ and B preserves the isomorphism between $\mathfrak{P}$ and A already set up in Theorem 31. Since $\mathfrak{P}$ is an ideal in $\mathfrak{P}^*$, it follows that A is an ideal in B; this statement can also be confirmed by virtue of the relations $(a, 0)(b, \beta) = (ab + a\beta + 0b, 0\beta) = (ab + a\beta, 0)$. In order to show that $\mathfrak{P}$ is not normal in $\mathfrak{P}^*$, we shall determine its orthocomplement relative to $\mathfrak{P}^*$. If $\mathfrak{a}$ is any semiprincipal ideal, or, more generally, a quite arbitrary ideal, we see that $\mathfrak{a}\mathfrak{a}(a) = \mathfrak{o}$ implies $\mathfrak{a}(a) = \mathfrak{a}\mathfrak{a}(a) = \mathfrak{o}$ and $a = 0$ whenever a is in $\mathfrak{a}$; and hence that, whenever $\mathfrak{a}\mathfrak{a}(a) = \mathfrak{o}$ for every a, the ideal $\mathfrak{a}$ coincides with $\mathfrak{o}$. Thus the orthocomplement of $\mathfrak{P}$ relative to $\mathfrak{P}^*$ or relative to $\mathfrak{S}$ consists of the ideal $\mathfrak{o}$ alone. Since $\mathfrak{P} \neq \mathfrak{P}^*$ and $\mathfrak{P} \neq \mathfrak{S}$ in the case under consideration, $\mathfrak{P}$ is not normal in $\mathfrak{P}^*$ or in $\mathfrak{S}$.

We turn now to the proof of the various numbered relations stated in the theorem. Of these certain ones have been established previously. Relations (1_1), (2_1), and (3_1) were proved in Theorem 31; (4_1) is part of Definition 8; and (4_2) follows at once from Theorem 23. By using the relation $\mathfrak{a}'(a) = \mathfrak{e} + \mathfrak{a}(a)$ which holds in $\mathfrak{S}$ by virtue of Theorem 30, we see at once that relations (3_2) and (3_3) follow from (3_1) and the known properties of the operation $+$. In a similar way, we deduce (2_2) from (2_1) and (3_1), writing

$$\mathfrak{a}(a)\mathfrak{a}'(b) = \mathfrak{a}(a)(\mathfrak{e} + \mathfrak{a}(b)) = \mathfrak{a}(a) + \mathfrak{a}(a)\mathfrak{a}(b) = \mathfrak{a}(a) + \mathfrak{a}(ab) = \mathfrak{a}(a + ab).$$

We obtain (1_1) from (2_2) by virtue of the relation $\mathfrak{a}(a) \vee \mathfrak{a}'(b) = (\mathfrak{a}'(a)\mathfrak{a}(b))'$. Similarly, we use (2_1) and (1_1) to establish (1_3) and (2_3) respectively with the help of the relations

$$\mathfrak{a}'(a) \vee \mathfrak{a}'(b) = (\mathfrak{a}(a)\mathfrak{a}(b))', \qquad \mathfrak{a}'(a)\mathfrak{a}'(b) = (\mathfrak{a}(a) \vee \mathfrak{a}(b))'.$$

With the results established in Theorems 23–32 we have covered the most important algebraic properties of the classes of ideals introduced in Definition 8. We call particular attention to the fact that Theorems 29–32 may be regarded as propositions concerning the imbedding of a Boolean ring A and its isomorph $\mathfrak{P}$ in Boolean rings of special type. Theorems 31 and 32 together give a new proof of the result of Theorem 1, as we have already observed. Theorems 30 and 31 give a similar imbedding theorem. Theorems 29 and 31 yield the result most interesting from the point of view of the present paper. They show that any Boolean ring A can be imbedded in a Boolean ring B which has a unit, which has the property that its normal ideals are all

principal, and which has under the unrestricted operations of (logical) addition and multiplication many of the formal properties of an algebra consisting of all the subclasses of a fixed abstract class under the operations of forming unrestricted unions and intersections; the Boolean ring B is isomorphic to the system of normal ideals in A. Evidently, B coincides with A when every normal ideal in A is principal, so that we can obtain nothing new when we apply this imbedding theorem to the Boolean ring B. If B had *all* the properties of an algebra of classes of the indicated type, in other words, if B were isomorphic to such an algebra, the representation problem would be solved by means of this imbedding theorem. In fact, this result fails to provide a solution; and, moreover, the considerations upon which it is based do not permit us to go beyond the Boolean ring B.†

4. **Prime ideals.** The ideals in a commutative ring may also be classified in terms of the relation of inclusion or divisibility. The three important types of ideal to be considered here are divisorless ideals, prime ideals, and primary ideals. An ideal $\mathfrak{a}$ is said to be divisorless if it is distinct from the ideal $\mathfrak{e}$ consisting of all the elements of the ring and has no ideal divisors other than $\mathfrak{a}$ and $\mathfrak{e}$. An ideal $\mathfrak{a}$ is said to be prime if it is distinct from $\mathfrak{e}$ and if, whenever it contains the product ab, it contains at least one of the factors a and b. Finally, an ideal $\mathfrak{a}$ is said to be primary if it is distinct from $\mathfrak{e}$ and if, whenever it contains the product ab but not the factor a, it contains some power, b^n, of the other factor b. One of the chief problems of ideal theory is the investigation of the properties of ideal products, the factors of which belong to one or another of the three classes just described. It is important to consider when an arbitrary ideal can be represented as such a product; when such a representation is unique; and when there exist divisorless, prime, or primary ideals at all. The answers to such questions serve to develop, and also to delimit, the generalization of the familiar arithmetic theory of prime numbers to abstract rings.

In the case of Boolean rings, we shall see that the arithmetic theory of prime ideals is equivalent to the theory of representations which is our major concern. It is therefore important that we study the properties of prime ideals in considerable detail. We shall begin with a series of conditions that an ideal be prime.

THEOREM 33. *The following assertions concerning an ideal* $\mathfrak{a}$ *in a Boolean ring* A *are equivalent:* $\mathfrak{a}$ *is divisorless,* $\mathfrak{a}$ *is prime,* $\mathfrak{a}$ *is primary.*

---

† Further light is shed on these remarks by the work of MacNeille, *The Theory of Partially Ordered Sets*, Harvard doctoral dissertation (1935), and Tarski, Fundamenta Mathematicae, vol. 24 (1935), pp. 177–198. See also the end of Chapter III, §3.

The last two assertions are obviously equivalent since $b^n = b$ in accordance with Definition 1. If $\mathfrak{a}$ is a divisorless ideal and a an element not in $\mathfrak{a}$, the ideal $\mathfrak{a} \vee \mathfrak{a}(a)$ must coincide with $\mathfrak{e}$ since it is a divisor of $\mathfrak{a}$ containing an element a not in $\mathfrak{a}$. Theorem 17 and the relation $\mathfrak{a} \vee \mathfrak{a}(a) = \mathfrak{e}$ show that an arbitrary element b can be expressed in the form $b = c \vee d$ where c and d are in $\mathfrak{a}$ and $\mathfrak{a}(a)$ respectively. Since $ad = d$, we have $d + c(a+d) = ad + ac + (ac)(ad) = ac \vee ad = ab$ and hence $d = ab + c(a+d)$, where $c(a+d)$ is in $\mathfrak{a}$. It therefore follows that, when ab is in $\mathfrak{a}$, the element d, and hence also the element $b = c \vee d$, is in $\mathfrak{a}$. Thus $\mathfrak{a}$ must contain b whenever it contains ab but not a; in other words, $\mathfrak{a}$ is a prime ideal. If $\mathfrak{a}$ is a prime ideal and $\mathfrak{b}$ an ideal divisor of $\mathfrak{a}$ not coincident with $\mathfrak{a}$, we select an element a in $\mathfrak{b}$ but not in $\mathfrak{a}$ and form the ideal $\mathfrak{a} \vee \mathfrak{a}(a)$; clearly we have $\mathfrak{a} \subset \mathfrak{a} \vee \mathfrak{a}(a) \subset \mathfrak{b}$. If b is an arbitrary element we write $b = (b+ab) + ab$ where $a(b+ab) = ab + ab = 0$ is in $\mathfrak{a}$ and ab is in $\mathfrak{a}(a)$. Since a is not in $\mathfrak{a}$ while the product $a(b+ab)$ is in $\mathfrak{a}$, the prime ideal $\mathfrak{a}$ must contain $b+ab$. It follows that b is in $\mathfrak{a} \vee \mathfrak{a}(a)$. Since b was arbitrary, we have $\mathfrak{a} \vee \mathfrak{a}(a) = \mathfrak{e}$ and hence $\mathfrak{b} = \mathfrak{e}$. Thus the ideal $\mathfrak{a}$ is divisorless, as we wished to prove.

In view of this theorem we need introduce only one symbol to denote the coincident classes of divisorless, prime, and primary ideals in a Boolean ring A; we shall use the letter $\mathfrak{E}$ for this purpose.

It is now convenient to restate the definition of a prime ideal in a Boolean ring in the following equivalent form, reposing in part upon Theorem 16:

THEOREM 34. *If the elements of a Boolean ring A be distributed between two non-void disjoint classes $\mathfrak{a}$ and $\mathfrak{b}$, then in order that $\mathfrak{a}$ be a prime ideal in A the following set of conditions is necessary and sufficient*:

(1) $a \in \mathfrak{a}$ *and* $b \in \mathfrak{a}$ *imply* $a \vee b \in \mathfrak{a}$;

(2) $a \in \mathfrak{a}$ *and* $b \in A$ *imply* $ab \in \mathfrak{a}$;

(3) $a \in \mathfrak{b}$ *and* $b \in \mathfrak{b}$ *imply* $ab \in \mathfrak{b}$.

We now have the following variation of this result:

THEOREM 35. *In Theorem* 34, *the condition* (2) *may be replaced by*

(2′) $a \in \mathfrak{b}$ *and* $b \in A$ *imply* $a \vee b \in \mathfrak{b}$.

We have to prove that (2) and (2′) are equivalent in the presence of (1) and (3). To prove (2′) from (1), (2), and (3), we proceed as follows: if a is in $\mathfrak{b}$ and b is arbitrary, the assumption that $a \vee b$ is in $\mathfrak{a}$ leads through (2) and the relation $a = a \vee ba = (a \vee b)a$ to the contradiction that a is in $\mathfrak{a}$. To prove (2) from (1), (2′), and (3), we argue as follows: if a is in $\mathfrak{a}$ and b is arbitrary, the assumption that ab is in $\mathfrak{b}$ leads through (2′) and the relation $a = a \vee ab$ to the contradiction that a is in $\mathfrak{b}$.

THEOREM 36. *If the Boolean ring A has a unit, the condition* (3) *of the set* (1), (2′), (3) *of Theorem* 35 *may be replaced by*†

(3′) $a \epsilon \mathfrak{b}$ *implies* $a' \epsilon \mathfrak{a}$.

The prime ideal $\mathfrak{a}$ *contains one, but not both, of the elements* a *and* a'.

We have to prove the equivalence of (3) and (3′) in the presence of (1) and (2′). To prove (3′) from (1), (2′), (3), we proceed as follows: if a is in $\mathfrak{b}$, then the assumption that a' is in $\mathfrak{b}$ leads through (3) and the relation $0 = a'a$ to the result that 0 is in $\mathfrak{b}$ and hence through (2′) and the relation $c = c \vee 0$ to the contradiction that $\mathfrak{b}$ contains every element c. To prove (3) from (1), (2′), (3′), we argue as follows: if a and b are in $\mathfrak{b}$, the assumption that ab is in $\mathfrak{a}$ leads through (1) and (3′) to the result that $a' \vee ab$ is in $\mathfrak{a}$ and hence through (2′) and the relation $a' \vee ab = a' \vee b$ to the contradiction that $a' \vee ab$ is also in $\mathfrak{b}$. Since the class $\mathfrak{a}$ is a prime ideal under conditions (1), (2′), (3′), it contains the product $a'a = 0$ and hence contains at least one of the factors a and a'. Since $\mathfrak{a}$ cannot contain $e = a' \vee a$ without coinciding with $\mathfrak{e}$ against hypothesis, we see that $\mathfrak{a}$ cannot contain both the elements a and a'.

As an application of the criteria given in Theorems 34–36 we have the following result:

THEOREM 37. *The Boolean ring B of Theorem* 1 *contains A as a prime ideal when A has no unit; and the system $\mathfrak{P}$ is a prime ideal in $\mathfrak{P}^*$ when $\mathfrak{P} \neq \mathfrak{P}^*$.*

In Theorem 32 we have already proved that A and $\mathfrak{P}$ are ideals. In order to show that they are prime ideals it is sufficient to establish condition (3) of Theorem 34. If we consider the Boolean ring of pairs (a, α) introduced in Theorem 1, we see that $(a, \epsilon)(b, \epsilon) = (ab + a + b, \epsilon)$ and hence that the elements of B which are not in A have the property demanded by (3). Similarly, the ideals in $\mathfrak{P}^*$ but not in $\mathfrak{P}$ have by virtue of the relation $\mathfrak{a}'(a)\mathfrak{a}'(b) = \mathfrak{a}'(a \vee b)$ of Theorem 32 (2) the property demanded by (3).

We next investigate the connections between the classification of the present section and that of §3. They are indicated in the following theorem:

THEOREM 38. *In a Boolean ring A, the classes $\mathfrak{E}$, $\mathfrak{N}$, and $\mathfrak{P}^*$ satisfy the inclusion relation $\mathfrak{E}\mathfrak{N} \subset \mathfrak{P}^*$. More precisely, an ideal $\mathfrak{p}$ is both prime and normal if and only if $\mathfrak{p} = \mathfrak{a}'(a)$ where a is an atomic element; and a prime ideal $\mathfrak{p}$ fails to be normal if and only if $\mathfrak{p}' = \mathfrak{o}$.*

† This theorem should be compared with recent work of Huntington, Bulletin of the American Mathematical Society, vol. 40 (1934), pp. 127–136 and pp. 137–142. It is easily seen that a Boolean ring with unit together with two subclasses $\mathfrak{a}$ and $\mathfrak{b}$ possessing properties (1), (2′), and (3′) is a system satisfying Huntington's postulates P1–P11 (on pp. 139–140); and, conversely, that any system satisfying those postulates is a Boolean ring with unit together with such subclasses $\mathfrak{a}$ and $\mathfrak{b}$. Huntington's postulates may therefore be regarded as postulates for a Boolean ring with unit together with a prime ideal.

If $\mathfrak{p}$ is any prime ideal, the fact that $\mathfrak{p}$ is divisorless taken together with the fact that $\mathfrak{p}'' \supset \mathfrak{p}$ shows that one of the relations $\mathfrak{p}'' = \mathfrak{p}$, $\mathfrak{p}'' = \mathfrak{e}$ must be true. If $\mathfrak{p}'' = \mathfrak{e}$, then $\mathfrak{p}$ is not normal and the relation $\mathfrak{p}' = \mathfrak{o}$ is valid. If $\mathfrak{p}'' = \mathfrak{p}$, then $\mathfrak{p}$ is normal and $\mathfrak{p}' \neq \mathfrak{o}$. When $\mathfrak{p}' \neq \mathfrak{o}$, we let a be an *arbitrary* element in $\mathfrak{p}'$ such that $a \neq 0$. Since $\mathfrak{p}' \supset \mathfrak{a}(a)$ and $\mathfrak{a}'(a) \neq \mathfrak{e}$, we see that $\mathfrak{p} = \mathfrak{p}'' \subset \mathfrak{a}'(a)$ and hence that $\mathfrak{p} = \mathfrak{a}'(a)$. This result shows that $\mathfrak{p}' = \mathfrak{a}(a)$ contains just one element a distinct from 0; in other words, that $b < a$ implies $b = 0$ or $b = a$. The element a is therefore atomic in accordance with Definition 3. To complete our proof, we must show that any ideal $\mathfrak{a}'(a)$, where a is an atomic element, is both prime and normal. Theorem 23 shows that $\mathfrak{a}'(a)$ is normal. To show that $\mathfrak{a}'(a)$ is prime we proceed as follows: if b and c are elements not belonging to $\mathfrak{a}'(a)$, we have $ab \neq 0$, $ac \neq 0$, and hence, by virtue of the fact that a is an atomic element, $ab = a$, $ac = a$; in consequence, we have $a(bc) = (ab)(ac) = aa = a \neq 0$ and conclude that bc is not in $\mathfrak{a}'(a)$; with the help of Theorem 34 (3), we thus find that $\mathfrak{p} = \mathfrak{a}'(a)$ is a prime ideal.

We shall now turn to the consideration of prime ideals in connection with operations upon ideals and the inclusion or divisibility relation between ideals. We first have

THEOREM 39. *If $\mathfrak{p}$ is a prime ideal in a Boolean ring A and $\mathfrak{a}$ is an arbitrary ideal, then*

(1) *the relations $\mathfrak{a} \subset \mathfrak{p}$, $\mathfrak{a}\mathfrak{p} = \mathfrak{p}$, $\mathfrak{a} \vee \mathfrak{p} = \mathfrak{p}$ are equivalent;*

(2) *the relations $\mathfrak{a} \not\subset \mathfrak{p}$, $\mathfrak{a}\mathfrak{p} \neq \mathfrak{p}$, $\mathfrak{a} \vee \mathfrak{p} \neq \mathfrak{p}$ are equivalent;*

(3) *one and only one of these two sets of mutually equivalent relations is valid. In case $\mathfrak{p}$ is normal, the relations $\mathfrak{a}\mathfrak{p}' = \mathfrak{o}$ and $\mathfrak{a}\mathfrak{p}' = \mathfrak{p}'$, where $\mathfrak{p}' \neq \mathfrak{o}$, are respectively equivalent to the relations in* (1) *and* (2) *respectively.*

We first prove the equivalence of the relations (1): $\mathfrak{a} \subset \mathfrak{p}$ obviously implies $\mathfrak{a}\mathfrak{p} = \mathfrak{a}$; $\mathfrak{a}\mathfrak{p} = \mathfrak{a}$ implies $\mathfrak{a} \vee \mathfrak{p} = \mathfrak{a}\mathfrak{p} \vee \mathfrak{p} = \mathfrak{p}$; and $\mathfrak{a} \vee \mathfrak{p} = \mathfrak{p}$ obviously implies $\mathfrak{a} \subset \mathfrak{p}$. We next observe that one and only one of the relations $\mathfrak{a} \vee \mathfrak{p} = \mathfrak{p}$, $\mathfrak{a} \vee \mathfrak{p} = \mathfrak{e}$ is valid, since $\mathfrak{p}$ is divisorless and $\mathfrak{a} \vee \mathfrak{p}$ is a divisor of $\mathfrak{p}$. Consequently, the relations (2) are simply the negations of the mutually equivalent relations in (1). Thus (2) and (3) both follow from (1). In case $\mathfrak{p}$ is normal, we know from Theorem 38 that $\mathfrak{p}$ is semiprincipal and that $\mathfrak{p}' \neq \mathfrak{o}$. Thus $\mathfrak{a}\mathfrak{p}' = \mathfrak{o}$ implies $\mathfrak{a} = \mathfrak{a}\mathfrak{e} = \mathfrak{a}(\mathfrak{p} \vee \mathfrak{p}') = \mathfrak{a}\mathfrak{p}$; and $\mathfrak{a}\mathfrak{p}' \neq \mathfrak{o}$ implies $\mathfrak{p} \subset \mathfrak{p} \vee \mathfrak{a}\mathfrak{p}' \neq \mathfrak{p}$, $\mathfrak{p} \vee \mathfrak{a}\mathfrak{p}' = \mathfrak{e}$, $\mathfrak{p} \vee \mathfrak{a} = \mathfrak{p} \vee \mathfrak{a}(\mathfrak{p} \vee \mathfrak{p}') = (\mathfrak{p} \vee \mathfrak{a}\mathfrak{p}) \vee \mathfrak{a}\mathfrak{p}' = \mathfrak{p} \vee \mathfrak{a}\mathfrak{p}' = \mathfrak{e}$, and $\mathfrak{a}\mathfrak{p}' = (\mathfrak{p} \vee \mathfrak{a}\mathfrak{p}')\mathfrak{p}' = \mathfrak{e}\mathfrak{p}' = \mathfrak{p}'$. The final statement of the theorem is thus evident.

THEOREM 40. *If $\mathfrak{p}$ is a prime ideal divisor of the ideal product $\mathfrak{a}\mathfrak{b}$ in a Boolean ring A, then $\mathfrak{p}$ is a divisor of at least one of the factors $\mathfrak{a}$ and $\mathfrak{b}$; in other words, $\mathfrak{p} \supset \mathfrak{a}\mathfrak{b}$ implies $\mathfrak{p} \supset \mathfrak{a}$ or $\mathfrak{p} \supset \mathfrak{b}$.*

If $\mathfrak{p}$ divides neither $\mathfrak{a}$ nor $\mathfrak{b}$, we have $\mathfrak{a} \vee \mathfrak{p} = \mathfrak{b} \vee \mathfrak{p} = \mathfrak{e}$, in accordance with the preceding theorem. By Theorem 18 (6), we see that the relation $\mathfrak{ab} \subset \mathfrak{p}$ leads to the absurd result $\mathfrak{p} = \mathfrak{ab} \vee \mathfrak{p} = (\mathfrak{a} \vee \mathfrak{p})(\mathfrak{b} \vee \mathfrak{p}) = \mathfrak{ee} = \mathfrak{e}$. Hence $\mathfrak{p}$ must divide $\mathfrak{a}$ or $\mathfrak{b}$. We point out that a corresponding result does not hold for infinite ideal products, as will be seen later. The failure to generalize Theorem 18 (6) to infinite products accounts for the distinction which we have to make here.

THEOREM 41. *If $\mathfrak{p}$ is a prime ideal in a Boolean ring A and $\mathfrak{a}$ is an arbitrary ideal, then at least one of the relations $\mathfrak{a} \subset \mathfrak{p}$, $\mathfrak{a}' \subset \mathfrak{p}$ is valid; if $\mathfrak{a}$ is simple, then only one is valid.*

Since $\mathfrak{p} \supset \mathfrak{o} = \mathfrak{a}'\mathfrak{a}$, at least one of the relations must hold by virtue of Theorem 40. If $\mathfrak{a}$ is simple and if both relations were to hold, we should have $\mathfrak{p} \supset \mathfrak{a} \vee \mathfrak{a}' = \mathfrak{e}$, $\mathfrak{p} = \mathfrak{e}$, against hypothesis.

In accordance with the introductory remarks of the present section and the result reached in Theorem 40, we may formulate for consideration the following statement:

FUNDAMENTAL PROPOSITION OF IDEAL ARITHMETIC. *In a Boolean ring A, every ideal other than $\mathfrak{e}$ is the product of all its prime ideal divisors.*

Our comments upon Theorem 40 indicate the possibility that an ideal may be expressible as the product merely of some, rather than of all, of its prime ideal divisors; and we shall see later that this possibility is realized in general. Now it is clear that the Fundamental Proposition of Ideal Arithmetic, and, indeed, all the preceding theorems concerning prime ideals, lack either meaning or content unless we can establish the following result:

FUNDAMENTAL EXISTENCE PROPOSITION. *In a Boolean ring A containing at least two elements, there exists at least one prime ideal.*

The exclusion of one-element Boolean rings is essential here since such a ring contains only the ideal $\mathfrak{o} = \mathfrak{e}$, which is not prime. We shall prove this proposition in Chapter IV, by means of transfinite methods; and we shall show that, if this proposition be true for every A, then the Fundamental Proposition of Ideal Arithmetic is true also for every A. The relation of these two central propositions to the theory of representations will be treated in Chapter IV.

5. **Congruences, ideals, and homomorphisms.** We must now consider the specialization of the general properties of ring-homomorphisms to the case of Boolean rings. We may first recall the principal results for arbitrary commutative rings. We denote a homomorphism from A to B by $A \rightarrow B$, referring to B as homomorphic to A or as a homomorph of A; and similarly we denote

an isomorphism from A to B by $A \longleftrightarrow B$, referring to B as isomorphic to A or as an isomorph of A. Any system with double composition homomorphic to a ring (or to a commutative ring) is itself a ring (or commutative ring). Here we confine our attention to the case of commutative rings. As we have previously remarked the determination of the homomorphs of an arbitrary algebraic system reduces essentially to the determination of all the congruences in the system. In the case of commutative rings, it is found that every congruence is a modular congruence. The modular congruences in a ring A are those defined in terms of the ideals of A by setting $a \equiv b$ if and only if $a-b$ is an element of a specified ideal $\mathfrak{a}$; in order to exhibit the ideal $\mathfrak{a}$ in our notation, we write $a \equiv b \pmod{\mathfrak{a}}$. It is easily verified that the modular congruences in a ring are congruences in the sense previously indicated. The proof that an arbitrary congruence is modular is easily given: the class $\mathfrak{a}$ of all elements a such that $a \equiv 0$ is an ideal since, if it contains a and b, it contains $a-b$ by virtue of the relations $a-b \equiv 0-0 \equiv 0$, and since, if it contains a, it contains ac by virtue of the relations $ac \equiv 0c \equiv 0$; and the relation $a \equiv b$ is equivalent to the relation $a-b \equiv 0$ and hence to the relation $a-b \in \mathfrak{a}$. If C is a congruence with the ideal $\mathfrak{a}$ as its modulus, we can introduce C as the fundamental equality in the given ring A so as to obtain the homomorph A^C, as previously indicated; we can also group the elements of A in classes of mutually congruent elements (mod $\mathfrak{a}$) and define the operations $+$ and $\cdot$ for these classes in the usual way so as to obtain the quotient ring $A/\mathfrak{a}$; and we see at once that A^C and $A/\mathfrak{a}$ are isomorphic. The chief theorem concerning the homomorphs of a ring A now asserts that every such homomorph is an isomorph of A^C and of $A/\mathfrak{a}$, where C is the congruence (mod $\mathfrak{a}$) and $\mathfrak{a}$ is the ideal consisting of all the elements of A which are carried into the zero element of the given homomorph.

We begin the consideration of Boolean rings with the following result:

THEOREM 42. *If the algebraic system B is homomorphic to a Boolean ring A with respect to the pair of operations $+$ and $\cdot$ or with respect to the pair of operations $\vee$ and $\cdot$, then B is homomorphic to A with respect to all three of the operations $+$, $\vee$, and $\cdot$; and B is a Boolean ring. If the algebraic system B is homomorphic to a Boolean ring A with unit with respect to the pair of operations $+$ and $\cdot$, with respect to the pair of operations $\vee$ and $'$, or with respect to the pair of operations $\vee$ and $\cdot$, then B is homomorphic to A with respect to all four of the operations $+$, $\vee$, $\cdot$, and $'$; and B is a Boolean ring with unit. The homomorphism $A \rightarrow B$ carries the zero element of A into the zero element of B, and the unit element of A, if any exists, into the unit element of B.*

The first statement of the theorem evidently refers to the two character-

izations of Boolean rings presented in Theorem 4. If B is a system with double composition homomorphic to a Boolean ring A under the operations $+$ and $\cdot$, then B is certainly a ring under the corresponding operations $+$ and $\cdot$ respectively. The law of idempotence is carried over from A to B by the given homomorphism, so that B is a Boolean ring. If the operation $\vee$ is now introduced in A and in B through the equation $a \vee b = a+b+ab$, the behavior of the homomorphism relative to the operations $+$ and $\cdot$ governs its behavior relative to the operation $\vee$; we thus see that the given homomorphism can be extended to cover the operation $\vee$ as well as the operations $+$ and $\cdot$, since the image in B of the element $a \vee b$ in A is evidently equal to the result of applying the operation $\vee$ to the images in B of the elements a and b respectively. On the other hand, if B is homomorphic to A under the operations $\vee$ and $\cdot$, we can show that the properties (1_1), (2_2), (3_1), (4_1), (5_1), (5_2), (6_1) and (6_2) hold in B as well as in A. All of these properties except (5_1) and (5_2) are identities in A and are therefore carried over by the homomorphism as identities in B. For example, if a^* is an arbitrary element of B and 0^* is the image in B of the zero element 0 in A, we can show that $a^* \vee 0^* = a^*$ in B: for there exists an element a in A which has a^* as its image in B, and the relation $a \vee 0 = a$ is carried by the given homomorphism into the relation to be proved. We may thus confine our attention to properties (5_1) and (5_2). If a^* and b^* are arbitrary elements in B satisfying the relation $b^*a^* = a^*$, we have to show that the system $x^* \vee a^* = b^*$, $x^*a^* = 0^*$ has a solution in B. There exist elements c and b in A which have a^* and b^* as their respective images in B under the given homomorphism. If we put $a = bc$, then $ba = b(bc) = bc = a$; and the image of a is $b^*a^* = a^*$. Consequently, the equations $x \vee a = b$, $xa = 0$ have a solution x. If we now take x^* as the image in B of the element x, we see that the equations satisfied by x in A are carried by the given homomorphism into the equations under consideration in B, and that x^* is the desired solution of these equations. Thus B is a generalized Boolean algebra and hence a Boolean ring in terms of the operations $+$ and $\cdot$, as shown in Theorem 4. We have to show that the image of the element $a+b$ in A is equal to the element a^*+b^* in B, where a^* and b^* are the images of a and b respectively. Since $a+b$ is the unique solution of the system $x \vee ab = a \vee b$, $x(ab) = 0$, we see that its image is a solution of the equations $x^* \vee a^*b^* = a^* \vee b^*$, $x^*(a^*b^*) = 0^*$; and since the latter equations have a^*+b^* as their unique solution in B, the desired result follows immediately. The homomorphism $A \rightarrow B$ can thus be extended to cover the operation $+$ as well as the operations $\vee$ and $\cdot$ with respect to which it was originally assumed to hold. The remaining statements of the theorem refer in a similar way to the characterizations of Boolean rings with unit presented in Theorems 2 and 4;

they can be established by similar, and slightly simpler, arguments. If A and B are Boolean rings, the homomorphism $A \to B$ carries the element 0 in A into an element 0^* in B which has the characteristic property of the zero element of B: for if a^* is any element of B and a an element of A with a^* as its image, the relation $a+0=a$ is carried by the homomorphism into the relation $a^*+0^*=a^*$. Similarly if A has a unit e, the homomorphism $A \to B$ carries e into an element e^* in B which has the characteristic property of a unit element in B: for if a^* is any element of B and a an element of A with a^* as its image, the relation $ea=a$ is carried by the homomorphism into the relation $e^*a^*=a^*$.

We next state the result which determines all the homomorphs of a Boolean ring; the proof has already been sketched above.

THEOREM 43. *In order that a Boolean ring B be homomorphic to a given Boolean ring A, it is necessary and sufficient that B be isomorphic to some quotient ring $A/\mathfrak{a}$, where $\mathfrak{a}$ is an ideal in A; in particular, the homomorphism $A \to B$ determines $\mathfrak{a}$ as the class of all elements in A which have the zero element in B as their image.*

We may consider also the characterization of all congruences in a Boolean ring. We have the following result:

THEOREM 44. *The only congruences in a Boolean ring A are the modular congruences. In order that $a \equiv b \pmod{\mathfrak{a}}$, where $\mathfrak{a}$ is an ideal in A, it is necessary and sufficient that $a+b$ belong to $\mathfrak{a}$, or that there exist elements c and d in $\mathfrak{a}$ for which $a \vee c = b \vee d$.*

That the only congruences in a commutative ring, and hence in a Boolean ring, are modular congruences, we have already proved in our introductory remarks. The relation $a+b=a-b$, which holds by virtue of Theorem 1, shows that $a \equiv b \pmod{\mathfrak{a}}$ if and only if $a+b$ belongs to $\mathfrak{a}$. Thus, if $a \equiv b \pmod{\mathfrak{a}}$, we have $a \vee (a+b) = a+(a+b)+a(a+b) = a+b+ab = a \vee b = b \vee (a+b)$ and hence $a \vee c = b \vee d$ where $c = d = a+b$ and $a+b$ is in $\mathfrak{a}$. On the other hand, if $a \vee c = b \vee d$ where c and d are in $\mathfrak{a}$, we have $a+(c+ac) = b+(d+bd)$ where $c+ac$ and $d+bd$ are in $\mathfrak{a}$ and hence $a+b = (c+ac)+(d+bd)$ where the element on the right is in $\mathfrak{a}$; and we conclude that $a \equiv b \pmod{\mathfrak{a}}$.

We shall next consider the behavior of subrings and ideals of a Boolean ring under an arbitrary homomorphism.

THEOREM 45. *If A_1 and A_2 are Boolean rings, if $\mathfrak{A}_1$ and $\mathfrak{A}_2$ are the classes of all subrings of A_1 and of A_2 respectively, and if the homomorphism $A_1 \to A_2$ determines the ideal $\mathfrak{a}_1$ in A, then the indicated homomorphism induces a homomorphism $\mathfrak{A}_1 \to \mathfrak{A}_2$ with respect to the operation of unrestricted addition. In par-*

ticular, the correspondences $A_1 \to A_2$, $\mathfrak{A}_1 \to \mathfrak{A}_2$ *have the following properties*:

(1) *if* $\mathfrak{b}_1$ *is a subring of* A_1, *the images of its elements under the homomorphism* $A_1 \to A_2$ *constitute a subring* $\mathfrak{b}_2$ *of* A_2 *corresponding to it under the homomorphism* $\mathfrak{A}_1 \to \mathfrak{A}_2$;

(2) *if* $\mathfrak{b}_1$ *is the subring generated by a non-void subclass* $\mathfrak{s}_1$ *of* A_1, *then its image* $\mathfrak{b}_2$ *under the homomorphisms* $A_1 \to A_2$ *and* $\mathfrak{A}_1 \to \mathfrak{A}_2$ *is the subring generated by the class* $\mathfrak{s}_2$ *of all images of elements of* $\mathfrak{s}_1$;

(3) *if* $\mathfrak{b}_2$ *is a subring of* A_2, *the class* $\mathfrak{b}_1$ *of all elements of* A_1 *with images in* $\mathfrak{b}_2$ *is a subring of* A_1 *with* $\mathfrak{b}_2$ *as its image under the homomorphisms* $A_1 \to A_2$ *and* $\mathfrak{A}_1 \to \mathfrak{A}_2$;

(4) *if* $\mathfrak{b}_1$ *and* $\mathfrak{c}_1$ *are subrings of* A_1 *with respective images* $\mathfrak{b}_2$ *and* $\mathfrak{c}_2$ *in* A_2, *then the relations* $\mathfrak{b}_1 \vee \mathfrak{a}_1 = \mathfrak{c}_1 \vee \mathfrak{a}_1$ *and* $\mathfrak{b}_2 = \mathfrak{c}_2$ *are equivalent.*

The proof is outlined essentially in (1), (2), and (3). The homomorphism $A_1 \to A_2$ takes the subring $\mathfrak{b}_1$ into a subclass $\mathfrak{b}_2$ of A_2. If a_2 and b_2 are in $\mathfrak{b}_2$, they are images respectively of elements a_1 and b_1 in $\mathfrak{b}_1$. It follows that a_2+b_2 and a_2b_2 are the images respectively of a_1+b_1 and a_1b_1 in $\mathfrak{b}_1$ and hence belong to $\mathfrak{b}_2$. Thus $\mathfrak{b}_2$ is a subring of A_2. The correspondence from $\mathfrak{A}_1$ to part of $\mathfrak{A}_2$ thus set up proves to be the desired homomorphism $\mathfrak{A}_1 \to \mathfrak{A}_2$. By the part of (1) now established, we see that the subring $\mathfrak{b}_1$ generated by a non-void subclass $\mathfrak{s}_1$ of A_1 is carried into a subring $\mathfrak{c}_2$ of A_2 containing the image of $\mathfrak{s}_1$, which we have denoted by $\mathfrak{s}_2$. Hence $\mathfrak{c}_2$ contains the subring $\mathfrak{b}_2$ generated by $\mathfrak{s}_2$. On the other hand, every element of $\mathfrak{c}_2$ is equal to a polynomial in terms of elements of $\mathfrak{s}_2$, by virtue of the construction of $\mathfrak{b}_1$ in terms of $\mathfrak{s}_1$. Hence $\mathfrak{c}_2$ is contained in $\mathfrak{b}_2$. We see therefore that $\mathfrak{b}_2 = \mathfrak{c}_2$, as we wished to prove. By the parts of (1) and (2) now established, we see that the class $\mathfrak{b}_1$ of all elements in A_1 with images in a given subring $\mathfrak{b}_2$ of A_2 generates a subring $\mathfrak{c}_1$ whose image in A_2 is the subring generated by $\mathfrak{b}_2$ and is thus $\mathfrak{b}_2$ itself. By definition, $\mathfrak{b}_1$ contains $\mathfrak{c}_1$. Hence $\mathfrak{b}_1$ coincides with $\mathfrak{c}_1$ and is therefore a subring of A_1 with $\mathfrak{b}_2$ as its image. By the part of (3) thus established, we see that the correspondence $\mathfrak{A}_1 \to \mathfrak{A}_2$ previously set up is a univocal correspondence taking the entire class $\mathfrak{A}_1$ into the entire class $\mathfrak{A}_2$. In view of the part of (2) already proved and in view of the definition of the sum of subrings as the least subring containing all the summands, we now see that the correspondence $\mathfrak{A}_1 \to \mathfrak{A}_2$ is a homomorphism with respect to the operation of unrestricted addition. All parts of the theorem, except (4), are thus proved. To establish (4), we first note that the image of the ideal $\mathfrak{a}_1$ under the correspondence $\mathfrak{A}_1 \to \mathfrak{A}_2$ is the ideal $\mathfrak{o}_2$ consisting of the zero element in A_2 alone. If $\mathfrak{b}_1$ and $\mathfrak{c}_1$ have the respective images $\mathfrak{b}_2$ and $\mathfrak{c}_2$ under the homomorphism $\mathfrak{A}_1 \to \mathfrak{A}_2$, then $\mathfrak{b}_1 \vee \mathfrak{a}_1$ and $\mathfrak{c}_1 \vee \mathfrak{a}_1$ have the respective images $\mathfrak{b}_2 \vee \mathfrak{o}_2 = \mathfrak{b}_2$, and $\mathfrak{c}_2 \vee \mathfrak{o}_2 = \mathfrak{c}_2$. It fol-

lows that $\mathfrak{b}_1 \vee \mathfrak{a}_1 = \mathfrak{c}_1 \vee \mathfrak{a}_1$ implies $\mathfrak{b}_2 = \mathfrak{c}_2$. Furthermore if a_1 is any element with image b_2 in $\mathfrak{b}_2$, then there exists an element b in $\mathfrak{b}_1$ which also has b_2 as its image. The fact that a_1 and b_1 have the same image implies that $a_1 \equiv b_1 \pmod{\mathfrak{a}_1}$, and hence that $a_1 = b_1 + (a_1 + b_1)$ belongs to $\mathfrak{b}_1 \vee \mathfrak{a}_1$. Thus $\mathfrak{b}_1 \vee \mathfrak{a}_1$ is the class of all elements in A_1 with images in $\mathfrak{b}_2$. Similarly, $\mathfrak{c}_1 \vee \mathfrak{a}_1$ is the class of all elements in A_1 with images in $\mathfrak{c}_2$. In consequence, $\mathfrak{b}_2 = \mathfrak{c}_2$ implies $\mathfrak{b}_1 \vee \mathfrak{a}_1 = \mathfrak{c}_1 \vee \mathfrak{a}_1$. This completes the proof of (4). Evidently the homomorphism $\mathfrak{A}_1 \to \mathfrak{A}_2$ determines a relation of congruence which is explicitly characterized by (4). We point out that the homomorphism $\mathfrak{A}_1 \to \mathfrak{A}_2$ cannot in general be extended to hold with respect to the multiplication of subrings. If $\mathfrak{a}_1$ is an ideal other than $\mathfrak{o}_1$ or $\mathfrak{e}_1$, we may determine elements b_1 and c_1 in A_1 such that $b_1 \neq c_1$, $b_1 \not\equiv 0_1 \pmod{\mathfrak{a}_1}$, $c_1 \not\equiv 0_1 \pmod{\mathfrak{a}_1}$, $b_1 \equiv c_1 \pmod{\mathfrak{a}_1}$. If A_2 is isomorphic to $A_1/\mathfrak{a}_1$, then the homomorphism $A_1 \to A_2$ takes b_1 and c_1 into the same element $b_2 \neq 0_2$. If now $\mathfrak{b}_1$ and $\mathfrak{c}_1$ are the subrings consisting of 0_1, b_1 and of 0_1, c_1 respectively, and if $\mathfrak{b}_2$ is the subring consisting of 0_2 and b_2, we see that $\mathfrak{b}_1\mathfrak{c}_1 = \mathfrak{o}_1$, $\mathfrak{b}_1 \to \mathfrak{b}_2$, $\mathfrak{c}_2 \to \mathfrak{b}_2$, $\mathfrak{b}_1\mathfrak{c}_1 \to \mathfrak{o}_2 \neq \mathfrak{b}_2 = \mathfrak{b}_2\mathfrak{b}_2$. This example indicates the impossibility of so extending the homomorphism.

THEOREM 46. *If A_1 and A_2 are Boolean rings, if $\mathfrak{J}_1$ and $\mathfrak{J}_2$ are the classes of all ideals in A_1 and in A_2 respectively, and if the homomorphism $A_1 \to A_2$ determines the ideal $\mathfrak{a}_1$ in A_1, then the indicated homomorphism induces a homomorphism $\mathfrak{J}_1 \to \mathfrak{J}_2$ with respect to the operations of unrestricted addition and finite multiplication. In particular, the correspondences $A_1 \to A_2$ and $\mathfrak{J}_1 \to \mathfrak{J}_2$ have properties* (1)–(4) *of Theorem* 45 *with the term "subring" everywhere replaced by the term "ideal."*

Since ideals are subrings and since the sum of ideals is an ideal in accordance with Theorem 18, we can prove all assertions of the present theorem, except (2) and those dealing with ideal multiplication, by showing that the correspondence $\mathfrak{A}_1 \to \mathfrak{A}_2$ takes $\mathfrak{J}_1$ into $\mathfrak{J}_2$ and then specializing Theorem 45 in the obvious way. Now, if $\mathfrak{b}_1$ is an ideal in A_1, its image $\mathfrak{b}_2$ in A_2 is a subring. Furthermore, if b_2 and c_2 are elements of $\mathfrak{b}_2$ and A_2 respectively, there exist elements b_1 and c_1 in $\mathfrak{b}_1$ and A_1 respectively with b_2 and c_2 as their respective images; and hence b_2c_2 is the image of b_1c_1, an element in $\mathfrak{b}_1$, and accordingly belongs to $\mathfrak{b}_2$. Thus $\mathfrak{b}_2$ is an ideal. On the other hand, if $\mathfrak{b}_2$ is an ideal in A_2, the class $\mathfrak{b}_1$ of all elements in A_1 with images in $\mathfrak{b}_2$ is a subring. Also, if b_1 and c_1 are in $\mathfrak{b}_1$ and A_1 respectively, their respective images b_2 and c_2 belong to $\mathfrak{b}_2$ and A_2 respectively; and hence b_1c_1 has the image b_2c_2 in $\mathfrak{b}_2$ and accordingly belongs to $\mathfrak{b}_1$. Thus $\mathfrak{b}_1$ is an ideal. The correspondence $\mathfrak{A}_1 \to \mathfrak{A}_2$ takes $\mathfrak{J}_1$ into $\mathfrak{J}_2$, as we wished to prove. We next discuss (2). The ideal $\mathfrak{a}_1(\mathfrak{s}_1)$ generated by a non-void subclass $\mathfrak{s}_1$ of A_1 has as its image an ideal $\mathfrak{b}_2$ which contains $\mathfrak{s}_2$, the

image of $\mathfrak{s}_1$. It follows that $\mathfrak{b}_2$ contains the ideal $\mathfrak{a}_2(\mathfrak{s}_2)$ generated by $\mathfrak{s}_2$. On the other hand, the class $\mathfrak{b}_1$ of all elements in A_1 with images in $\mathfrak{a}_2(\mathfrak{s}_2)$ is an ideal which contains $\mathfrak{s}_1$ and hence also $\mathfrak{a}_1(\mathfrak{s}_1)$. It follows that $\mathfrak{a}_2(\mathfrak{s}_2)$ contains $\mathfrak{b}_2$. We therefore conclude that $\mathfrak{a}_2(\mathfrak{s}_2)$ is the image of $\mathfrak{a}_1(\mathfrak{s}_1)$, as we wished to prove. We still have to establish the assertion that the correspondence $\mathfrak{I}_1 \rightarrow \mathfrak{I}_2$ is a homomorphism with respect to finite multiplication. Since it is evident that the image of the product $\mathfrak{b}_1\mathfrak{c}_1$ of ideals $\mathfrak{b}_1$ and $\mathfrak{c}_1$ in A_1 is an ideal contained in the images $\mathfrak{b}_2$ and $\mathfrak{c}_2$ of $\mathfrak{b}_1$ and $\mathfrak{c}_1$ respectively, we see that the image is contained also in the product $\mathfrak{b}_2\mathfrak{c}_2$. On the other hand, if a_2 is an arbitrary element of $\mathfrak{b}_2\mathfrak{c}_2$, it is expressible in the form $a_2 = b_2c_2$, where b_2 is in $\mathfrak{b}_2$ and c_2 in $\mathfrak{c}_2$, as we have proved in Theorem 18; and the elements b_2 and c_2 are the images of elements b_1 and c_1 respectively in $\mathfrak{b}_1$ and $\mathfrak{c}_1$ respectively. Thus a_2 is the image of the element b_1c_1 in $\mathfrak{b}_1\mathfrak{c}_1$ and hence belongs to the image of $\mathfrak{b}_1\mathfrak{c}_1$. We conclude therefore that the image of $\mathfrak{b}_1\mathfrak{c}_1$ coincides with the product $\mathfrak{b}_2\mathfrak{c}_2$ of the images of $\mathfrak{b}_1$ and $\mathfrak{c}_1$ respectively, and thereby bring the proof to a close.

We may point out that the homomorphism $\mathfrak{I}_1 \rightarrow \mathfrak{I}_2$ cannot be extended to the operation of orthocomplementation. This we shall show by examples cited in connection with the following theorem.

Our next result deals with the behavior of ideals relative to the classification of §3. We have

THEOREM 47. *If A_1 and A_2 are Boolean rings, the homomorphism $A_1 \rightarrow A_2$ carries every principal (semiprincipal, simple) ideal in A_1 into a principal (semiprincipal, simple) ideal in A_2; but it may carry a normal ideal in A_1 into a non-normal ideal in A_2.*

If $\mathfrak{b}_1$ is a principal ideal in A_1, its image in A_2 is an ideal which, considered as a Boolean ring, is homomorphic to $\mathfrak{b}_1$ and hence has a unit in accordance with Theorem 42. It follows that $\mathfrak{b}_2$ is a principal ideal in A_2. We pass next to the case of a simple ideal $\mathfrak{b}_1$ in A_1. The images of $\mathfrak{b}_1$ and $\mathfrak{b}_1'$ in A_2 are ideals $\mathfrak{b}_2$ and $\mathfrak{c}_2$ respectively. The relation $\mathfrak{b}_1\mathfrak{b}_1' = \mathfrak{o}_1$ implies the relation $\mathfrak{b}_2\mathfrak{c}_2 = \mathfrak{o}_2$ in accordance with Theorem 46 and hence the relation $\mathfrak{c}_2 \subset \mathfrak{b}_2'$. The relation $\mathfrak{b}_1 \vee \mathfrak{b}_1' = \mathfrak{e}_1$ similarly implies the relation $\mathfrak{b}_2 \vee \mathfrak{c}_2 = \mathfrak{e}_2$. Thus we have $\mathfrak{b}_2 \vee \mathfrak{b}_2' \supset \mathfrak{b}_2 \vee \mathfrak{c}_2 = \mathfrak{e}_2$, $\mathfrak{b}_2 \vee \mathfrak{b}_2' = \mathfrak{e}_2$. It follows that $\mathfrak{b}_2$ is a simple ideal. In case $\mathfrak{b}_1$ is semiprincipal but not principal, $\mathfrak{b}_1'$ is principal. The image $\mathfrak{b}_2$ of $\mathfrak{b}_1$ is simple by the result just established; and the image $\mathfrak{c}_2$ of $\mathfrak{b}_1'$ is principal. Thus $\mathfrak{b}_2' = \mathfrak{b}_2'\mathfrak{e}_2 = \mathfrak{b}_2'(\mathfrak{b}_2 \vee \mathfrak{c}_2) = \mathfrak{b}_2'\mathfrak{c}_2 = \mathfrak{c}_2$ by virtue of the relations noted above. Since $\mathfrak{b}_2$ is simple and $\mathfrak{b}_2'$ is principal, it follows that $\mathfrak{b}_2 = \mathfrak{b}_2''$ is semiprincipal, as we wished to prove. In order to show that similar results do not hold in the case of normal ideals, let us consider the case where the homomorphism $A_1 \rightarrow A_2$ determines an ideal $\mathfrak{a}_1$ which is normal but not simple, with a view to

showing that the normal ideal $\mathfrak{a}_1'$ is carried into a non-normal ideal $\mathfrak{b}_2$. The homomorphism carries $\mathfrak{a}_1$ into $\mathfrak{o}_2$, $\mathfrak{a}_1 \vee \mathfrak{a}_1'$ into $\mathfrak{b}_2$. We shall show that $\mathfrak{b}_2 \neq \mathfrak{e}_2$, $\mathfrak{b}_2' = \mathfrak{o}_2$, so that the relation $\mathfrak{b}_2'' = \mathfrak{b}_2$ is impossible. It is then clear that $\mathfrak{b}_2$ is not normal. Since $\mathfrak{a}_1 \vee \mathfrak{a}_1' \neq \mathfrak{e}_1$, we see with the aid of Theorem 46 (4) that $\mathfrak{b}_2 \neq \mathfrak{e}_2$. Now let $\mathfrak{c}_1$ be the class of all elements in A_1 with images in $\mathfrak{b}_2'$. By virtue of Theorem 46, the image of the ideal $\mathfrak{c}_1(\mathfrak{a}_1 \vee \mathfrak{a}_1')$ is $\mathfrak{b}_2'(\mathfrak{o}_2 \vee \mathfrak{b}_2) = \mathfrak{o}_2$. Hence we have $\mathfrak{c}_1(\mathfrak{a}_1 \vee \mathfrak{a}_1') \subset \mathfrak{a}_1$, $\mathfrak{c}_1\mathfrak{a}_1' = \mathfrak{c}_1(\mathfrak{a}_1 \vee \mathfrak{a}_1')\mathfrak{a}_1' \subset \mathfrak{a}_1\mathfrak{a}_1' = \mathfrak{o}_1$, $\mathfrak{c}_1\mathfrak{a}_1' = \mathfrak{o}_1$, $\mathfrak{c}_1 \subset \mathfrak{a}_1'' = \mathfrak{a}_1$, and $\mathfrak{b}_2' = \mathfrak{o}_2$, as we wished to show. The fact that the homomorphism $A_1 \rightarrow A_2$ does not always carry normal ideals into normal ideals shows that the induced homomorphism $\mathfrak{I}_1 \rightarrow \mathfrak{I}_2$ does not always hold relative to the operation of orthocomplementation, since an ideal $\mathfrak{b}_1$ in A_1 may satisfy the relation $\mathfrak{b}_1'' = \mathfrak{b}_1$ and yet have an image $\mathfrak{b}_2$ for which $\mathfrak{b}_2'' \neq \mathfrak{b}_2$.

We shall also consider the behavior of prime ideals under a homomorphism, obtaining the following result:

THEOREM 48. *If A_1 and A_2 are Boolean rings and if $\mathfrak{a}_1$ is the ideal determined in A_1 by the homomorphism $A_1 \rightarrow A_2$, then the indicated homomorphism carries a prime ideal $\mathfrak{p}_1$ in A_1 into an ideal $\mathfrak{p}_2$ in A_2 which is prime or coincides with $\mathfrak{e}_2$ according as $\mathfrak{p}_1$ contains $\mathfrak{a}_1$ or not. If $\mathfrak{p}_2$ is a prime ideal in A_2, the class $\mathfrak{p}_1$ of all elements in A_1 with images in $\mathfrak{p}_2$ is a prime ideal in A_1.*

If $\mathfrak{p}_1$ is prime, then $\mathfrak{p}_1 \vee \mathfrak{a}_1 = \mathfrak{p}_1$ or $\mathfrak{p}_1 \vee \mathfrak{a}_1 = \mathfrak{e}_1$, according as $\mathfrak{p}_1$ contains $\mathfrak{a}_1$ or not, as we have proved in Theorem 39. Since $\mathfrak{p}_1$ and $\mathfrak{p}_1 \vee \mathfrak{a}_1$ have a common image $\mathfrak{p}_2$ in A_2, we see that $\mathfrak{p}_2 \neq \mathfrak{e}_2$ or $\mathfrak{p}_2 = \mathfrak{e}_2$ according as $\mathfrak{p}_1$ contains $\mathfrak{a}_1$ or not. In case $\mathfrak{p}_2 \neq \mathfrak{e}_2$, let $\mathfrak{b}_2$ be an ideal which contains $\mathfrak{p}_2$, and let $\mathfrak{b}_1$ be the ideal consisting of all elements in A_1 with images in $\mathfrak{b}_2$. Since $\mathfrak{b}_1$ evidently contains $\mathfrak{p}_1$, we must have either $\mathfrak{b}_1 = \mathfrak{p}_1$ or $\mathfrak{b}_1 = \mathfrak{e}_1$ and hence either $\mathfrak{b}_2 = \mathfrak{p}_2$ or $\mathfrak{b}_2 = \mathfrak{e}_2$. It follows that $\mathfrak{p}_2$ is divisorless and therefore prime, by virtue of Theorem 33. If $\mathfrak{p}_2$ is a prime ideal in A_2, the ideal $\mathfrak{p}_1$ consisting of all elements in A_1 with images in $\mathfrak{p}_2$ evidently contains $\mathfrak{a}_1$. Since $\mathfrak{p}_2 \neq \mathfrak{e}_2$, we have $\mathfrak{p}_1 \neq \mathfrak{e}_1$ likewise. If $\mathfrak{b}_1$ is an ideal containing $\mathfrak{p}_1$, its image $\mathfrak{b}_2$ contains $\mathfrak{p}_2$ and must therefore coincide with $\mathfrak{p}_2$ or with $\mathfrak{e}_2$. When $\mathfrak{b}_2 = \mathfrak{p}_2$, we must have $\mathfrak{b}_1 \vee \mathfrak{a}_1 = \mathfrak{p}_1 \vee \mathfrak{a}_1$ by Theorem 46 (4); but since $\mathfrak{b}_1 \supset \mathfrak{p}_1 \supset \mathfrak{a}_1$, this relation implies $\mathfrak{b}_1 = \mathfrak{p}_1$. Similarly, when $\mathfrak{b}_2 = \mathfrak{e}_2$, we must have $\mathfrak{b}_1 \vee \mathfrak{a}_1 = \mathfrak{e}_1 \vee \mathfrak{a}_1$ and hence $\mathfrak{b}_1 = \mathfrak{e}_1$. It follows that $\mathfrak{p}_1$ is divisorless and therefore prime.

The final theorem concerning homomorphisms offers a new criterion to determine whether an ideal is prime or not.

THEOREM 49. *In order that an ideal $\mathfrak{p}$ in a Boolean ring A be prime, it is necessary and sufficient that $A/\mathfrak{p}$ be a two-element Boolean ring.*

If $A/\mathfrak{p}$ is a two-element Boolean ring, it consists of a zero element 0^* and

a unit element e^*. An element a in A belongs to $\mathfrak{p}$ or not according as its image under the homomorphism $A \to A/\mathfrak{p}$ is 0^* or e^*. The homomorphism cannot take the product ab into 0^* if it takes a and b both into e^*; in other words, if $\mathfrak{p}$ contains ab, it contains at least one of the elements a and b, and is therefore prime. If $\mathfrak{p}$ is prime, then $A/\mathfrak{p}$ contains no divisors of zero. For if a^* and b^* are elements of $A/\mathfrak{p}$ such that $a^*b^* = 0^*$, they are images of elements a and b in A such that ab is in $\mathfrak{p}$; since $\mathfrak{p}$ is prime, at least one of these elements is in $\mathfrak{p}$, and at least one of the elements a^* and b^* is equal to 0^*. By Theorem 1, the Boolean ring $A/\mathfrak{p}$ cannot have more than two elements. On the other hand, the relation $\mathfrak{p} \neq \mathfrak{e}$ shows that $A/\mathfrak{p}$ has at least two elements. The proof is thus completed.

6. **Direct sums.** In the case of Boolean rings, the theory of representations has a very close connection with the concept of the direct sum and the problem of representing a given Boolean ring as a direct sum. We shall therefor close the present chapter with a brief section devoted to this topic. The direct sum of Boolean rings A_α where the index α ranges over an arbitrary finite or infinite class A is the algebraic system $S_{\alpha \in A} A_\alpha$ described as follows: the elements of the system are all the functions f defined over A with values $f(\alpha)$ in A_α; the relation of equality in the system is defined by putting $f = g$ if and only if $f(\alpha) = g(\alpha)$ for every α; the fundamental operations of the system are addition and multiplication, the sum $f+g$ being defined as the function h such that $h(\alpha) = f(\alpha) + g(\alpha)$ in A_α for every α and the product fg being defined as the function k such that $k(\alpha) = f(\alpha)g(\alpha)$ in A_α for every α. In case the class A consists of the integers 1 and 2, we denote the direct sum of the Boolean rings A_α by $A_1 \vee A_2$; this direct sum may evidently be described as the class of all ordered pairs (a_1, a_2) where a_1 and a_2 are elements of A_1 and of A_2 respectively, with $(a_1, a_2) = (b_1, b_2)$ if and only if $a_1 = b_1$ and $a_2 = b_2$ and with $(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 + b_2)$, $(a_1, a_2)(b_1, b_2) = (a_1 b_1, a_2 b_2)$. The following properties of the direct sum may be stated without formal proof:

THEOREM 50. *The direct sum $S_{\alpha \in A} A_\alpha$ of Boolean rings A_α where α ranges over the class A is a Boolean ring. Its zero element is the function 0 which for each α has as its value $0(\alpha)$ the zero element of A_α. It has a unit if and only if every A_α has a unit; if A_α has a unit e_α for every α, then the function e which for each α has the value $e(\alpha) = e_\alpha$ is the unit of the direct sum.*

It is an important problem to determine when a Boolean ring is isomorphic to a direct sum of Boolean rings in a non-trivial manner. Since the direct sum of a Boolean ring A with a one-element Boolean ring is isomorphic to A, we must evidently exclude one-element summands as trivial. A Boolean ring which is representable as the direct sum of two or more Boolean rings, each

with at least two elements, is said to be reducible; and a Boolean ring which is not so representable is said to be irreducible. The reducibility of a Boolean ring is correlated with its ideal structure by virtue of the following result:

THEOREM 51. *If* $\mathfrak{a}$ *is a simple ideal in a Boolean ring* A, *then* $A \longleftrightarrow (A/\mathfrak{a}) \vee (A/\mathfrak{a}')$, $\mathfrak{a} \longleftrightarrow A/\mathfrak{a}'$, $\mathfrak{a}' \longleftrightarrow A/\mathfrak{a}$. *Conversely, if* A, A_1, *and* A_2 *are Boolean rings such that* $A \longleftrightarrow A_1 \vee A_2$, *then there exists a simple ideal* $\mathfrak{a}$ *in* A *such that* $A_1 \longleftrightarrow A/\mathfrak{a}$, $A_2 \longleftrightarrow A/\mathfrak{a}'$.

If $\mathfrak{a}$ is a simple ideal in A, then every element a in A can be represented in a unique manner as the sum a_1+a_2 of elements a_1 and a_2 in $\mathfrak{a}$ and $\mathfrak{a}'$ respectively. To establish the representation, we note that $\mathfrak{a}$ and $\mathfrak{a}'$ are simple by Theorem 30 and hence that $\mathfrak{a}(a)\mathfrak{a}$ and $\mathfrak{a}(a)\mathfrak{a}'$ are principal ideals $\mathfrak{a}(a_1)$ and $\mathfrak{a}(a_2)$ by Theorem 26. The relations $\mathfrak{a}(a_1) \subset \mathfrak{a}$, $\mathfrak{a}(a_2) \subset \mathfrak{a}'$, $\mathfrak{a}(a_1+a_2) = \mathfrak{a}(a_1)+\mathfrak{a}(a_2) = \mathfrak{a}(a)\mathfrak{a}+\mathfrak{a}(a)\mathfrak{a}' = \mathfrak{a}(a)(\mathfrak{a}+\mathfrak{a}') = \mathfrak{a}(a)(\mathfrak{a}+\mathfrak{a}'+\mathfrak{a}\mathfrak{a}') = \mathfrak{a}(a)(\mathfrak{a} \vee \mathfrak{a}') = \mathfrak{a}(a)\mathfrak{e} = \mathfrak{a}(a)$ show that a_1 is in $\mathfrak{a}$, that a_2 is in $\mathfrak{a}'$, and that $a_1+a_2=a$. If a has a second representation $a=b_1+b_2$ where b_1 and b_2 are in $\mathfrak{a}$ and $\mathfrak{a}'$ respectively, the relation $a_1+a_2=b_1+b_2$ implies the relation $a_1+b_1=a_2+b_2$, which shows that a_1+b_1 and a_2+b_2 belong both to $\mathfrak{a}$ and to $\mathfrak{a}'$. Hence we have $a_1+b_1=a_2+b_2=0$, $a_1=b_1$, $a_2=b_2$. The representation is therefore unique. In terms of this representation we set up a biunivocal correspondence from A to the class of ordered pairs (a_1, a_2) where a_1 is in $\mathfrak{a}$ and a_2 in $\mathfrak{a}'$: we put $a \longleftrightarrow (a_1, a_2)$ if and only if $a=a_1+a_2$. It is evident that this correspondence has the property that $a \longleftrightarrow (a_1, a_2)$ and $b \longleftrightarrow (b_1, b_2)$ imply $a+b \longleftrightarrow (a_1+b_1, a_2+b_2)$, $ab \longleftrightarrow (a_1b_1, a_2b_2)$. Consequently, we see that, if we regard $\mathfrak{a}$ and $\mathfrak{a}'$ as Boolean rings, their direct sum is isomorphic to A. Furthermore, if $a=a_1+a_2$ and $b=b_1+b_2$, we see that $a \equiv a_2 \pmod{\mathfrak{a}}$, $b \equiv b_2 \pmod{\mathfrak{a}}$, $a+b \equiv a_2+b_2 \pmod{\mathfrak{a}}$, and $ab \equiv a_2b_2 \pmod{\mathfrak{a}}$. We conclude at once that, if we regard $\mathfrak{a}'$ as a Boolean ring, then $\mathfrak{a}'$ is isomorphic to $A/\mathfrak{a}$. Since $\mathfrak{a}'$ is simple and $(\mathfrak{a}')' = \mathfrak{a}$, the isomorphism $\mathfrak{a} \longleftrightarrow A/\mathfrak{a}'$ is also valid. The fact that A is isomorphic to the direct sum of $\mathfrak{a}$ and $\mathfrak{a}'$ shows that A is also isomorphic to the direct sum of their isomorphs $A/\mathfrak{a}$ and $A/\mathfrak{a}'$, taken in inverted order.

If $A \longleftrightarrow A_1 \vee A_2$, we see that the class $\mathfrak{a}_1$ of all pairs $(a_1, 0_2)$ in $A_1 \vee A_2$ is an ideal; and similarly that the class $\mathfrak{a}_2$ of all pairs $(0_1, a_2)$ is an ideal. If the ideals $\mathfrak{a}_1$ and $\mathfrak{a}_2$ are regarded as Boolean rings they are isomorphic to A_1 and to A_2 respectively. It is furthermore evident that $\mathfrak{a}_1\mathfrak{a}_2$ consists of the zero element $(0_1, 0_2)$ in $A_1 \vee A_2$ and that $\mathfrak{a}_1 \vee \mathfrak{a}_2$ consists of all elements (a_1, a_2) in $A_1 \vee A_2$. Thus if we pass to the Boolean ring A by means of the isomorphism $A \longleftrightarrow A_1 \vee A_2$, we see that A contains ideals $\mathfrak{b}$ and $\mathfrak{a}$ isomorphic to $\mathfrak{a}_1$ and $\mathfrak{a}_2$ respectively and hence also to A_1 and A_2 respectively, these ideals satisfying the relations $\mathfrak{a}\mathfrak{b}=\mathfrak{o}$, $\mathfrak{a} \vee \mathfrak{b} = \mathfrak{e}$. Since $\mathfrak{a}\mathfrak{b}=\mathfrak{o}$ implies $\mathfrak{a}' \supset \mathfrak{b}$, we see that

$\mathfrak{a} \vee \mathfrak{a}' \supset \mathfrak{a} \vee \mathfrak{b} = \mathfrak{e}$ and hence that $\mathfrak{a}$ is simple. Similarly $\mathfrak{b}$ is simple. Since $\mathfrak{a}' = \mathfrak{a}'\mathfrak{e} = \mathfrak{a}'(\mathfrak{a} \vee \mathfrak{b}) = \mathfrak{a}'\mathfrak{b}$, we have $\mathfrak{a}' \subset \mathfrak{b}$ and hence $\mathfrak{a}' = \mathfrak{b}$, $\mathfrak{b}' = \mathfrak{a}$. The isomorphisms $A_1 \longleftrightarrow \mathfrak{b}$, $\mathfrak{b} \longleftrightarrow A/\mathfrak{b}'$ thus lead to the isomorphism $A_1 \longleftrightarrow A/\mathfrak{a}$. The isomorphisms $A_2 \longleftrightarrow \mathfrak{a}$, $\mathfrak{a} \longleftrightarrow A/\mathfrak{a}'$ lead similarly to the isomorphism $A_2 \longleftrightarrow A/\mathfrak{a}'$.

An immediate consequence of Theorem 51 is the following proposition about reducibility:

THEOREM 52. *A Boolean ring A is reducible if and only if it has more than two elements.*

From the preceding theorem we see that A is reducible if and only if the class $\mathfrak{S}$ of all simple ideals contains an ideal $\mathfrak{a}$ other than $\mathfrak{o}$ or $\mathfrak{e}$, the relations $\mathfrak{a}' \neq \mathfrak{e}$, $\mathfrak{a}' \neq \mathfrak{o}$ being equivalent to the relations $\mathfrak{a} \neq \mathfrak{o}$, $\mathfrak{a} \neq \mathfrak{e}$ respectively. Since the class $\mathfrak{P}$ of all principal ideals is contained in $\mathfrak{S}$ and has the same cardinal number as A, we see that, whenever A contains more than two elements, $\mathfrak{S}$ contains more than the two ideals $\mathfrak{o}$ and $\mathfrak{e}$ and A is reducible. If A has one or two elements, then A has a unit by Theorem 1 and $\mathfrak{S}$ coincides with $\mathfrak{P}$ by Theorem 25. Thus $\mathfrak{S}$ contains only the ideals $\mathfrak{o}$ and $\mathfrak{e}$, and A is irreducible.

We conclude with the consideration of the direct sums of two-element Boolean rings, that is, of non-trivial irreducible Boolean rings. We have

THEOREM 53. *A Boolean ring A is isomorphic to the direct sum of two-element Boolean rings A_α, where α ranges over the class* A, *if and only if it is isomorphic to the algebra of all subclasses of* A.

It is evidently sufficient for us to prove that the direct sum of *identical* two-element Boolean rings A_α, where α is in A, is isomorphic to the algebra of all subclasses of A. To a function f in the direct sum we order the class A_f of all elements α such that $f(\alpha) = e$. It is then evident that $\mathrm{A}_f = \mathrm{A}_g$ if and only if $f = g$, that $\mathrm{A}_{f+g} = \mathrm{A}_f \Delta \mathrm{A}_g$, and that $\mathrm{A}_{fg} = \mathrm{A}_f \mathrm{A}_g$. It follows that the correspondence $f \longleftrightarrow \mathrm{A}_f$ determines an isomorphism between A and an algebra of subclasses of A. That the latter algebra is the algebra of *all* subclasses of A we see as follows: if B is an arbitrary subclass of A, the function f which is equal to e or to 0 according as α is in B or not is an element of the direct sum and has the property that $\mathrm{A}_f = \mathrm{B}$.

This theorem is significant from two points of view. In the first place, it shows that, although every Boolean ring of more than two elements is reducible, such a ring need not be completely reducible in the sense that it is representable as the direct sum of irreducible summands or components. The complete reduction of finite Boolean rings is possible in accordance with Theorem 12; but that of infinite Boolean rings is in general impossible. In the second place, the theorem shows the existence of a connection between

the theory of direct sums and that of the representation of Boolean rings by algebras of classes.

CHAPTER III. ALGEBRAS OF CLASSES

1. **The construction of algebras of classes.** By an algebra of classes or concrete Boolean ring we shall mean a Boolean ring which has as elements certain subclasses of a fixed finite or infinite class and as operations those of forming the union (modulo 2) (or symmetric difference) and the intersection of classes. It is easily verified that the class of all subclasses of a fixed class is such a concrete Boolean ring, in accordance with Definition 1. Furthermore, this algebra of classes contains as a subring every concrete Boolean ring with elements which are subclasses of the same fixed class. The object of the present chapter is to discuss the structure of concrete Boolean rings in terms of the general theory developed in the preceding pages.

While the explicit construction of algebras of classes is not of primary concern to us here, some brief remarks upon this subject will permit us to review briefly a few of the points of contact between the present abstract theory and other branches of mathematics, and also to illustrate some of the concepts which we have had occasion to introduce. The methods available for constructing concrete Boolean rings on the basis of an entirely abstract class are essentially those indicated in Chapter II, §§1–2. To them we may add the use of cardinal numbers in forming rings and ideals: thus the countable subclasses of a non-countable class, the finite subclasses of an infinite class afford examples of Boolean rings which are ideals in the algebras of all subclasses of the respective basic classes. Since the union of classes corresponds in the case of a concrete Boolean ring to the abstract element $a \vee b$, we have been at some pains, particularly in Theorems 14, 16, and 17, to show how the operation $\vee$ can be used in place of the operation $+$ in carrying out such constructions. It is only in the case of a class in which there are given relations, operations, and properties other than those of the pure logic of classes that other methods of construction can be applied to obtain algebras of classes. For example, in the theory of measure in euclidean space of n dimensions, it is the combination of the geometrical and topological properties of the space with the operations of the logic of classes which leads to the study of the various concrete Boolean rings consisting respectively of the subclasses described as follows: Lebesgue measurable sets, Borel measurable sets, Lebesgue measurable sets of finite measure, Borel measurable sets of finite measure, and sets of zero measure. It will be noted that the last-named algebra of classes is an ideal in the algebra of all subsets of the given space. Again, in general topology, analogous circumstances lead to the study of

Boolean rings consisting of the subclasses variously described as follows: Borel sets, nowhere dense sets, and sets of the first category. It will be noted that the two last named algebras of classes are ideals in the algebra of all subclasses of the basic topological space. Similar ideals are of such frequent occurrence and of such real importance in topology that they have been given the special designation "additive hereditary classes" by Kuratowski.†

We may point out that some of the algebras of classes cited above are Boolean rings without unit. For instance, the finite subclasses of an infinite class and the Lebesgue measurable sets of finite measure constitute such rings. It is likewise possible to use these and similar concrete Boolean rings to illustrate other points raised in the development of the general theory. At the end of the present chapter we shall discuss the Boolean ring of Lebesgue measurable sets of finite measure in sufficient detail to uncover a number of such illustrative properties.

2. **Reduction and equivalence.** Before proceeding with the study of concrete Boolean rings, we shall describe a canonical form for such rings, and show how any algebra of classes can be reduced to canonical form; that is, how it can be replaced by an isomorphic algebra of classes in that form.

DEFINITION 10. *If A is a Boolean ring with elements $a, b, c, \cdots$ which are subclasses of a fixed class $\mathrm{E}=\mathrm{E}(A)$ with elements $\alpha, \beta, \gamma, \cdots$, then A is said to be a reduced algebra of classes when it has the following property: every elemen α in E is contained in some element of A and is the only element of E common to all the elements of A containing it.*

We can now prove the following result:

THEOREM 54. *Every algebra of classes with more than one element is isomorphic to a reduced algebra of classes, by virtue of an element-to-element correspondence of the basic classes.*

The case of a one-element algebra of classes is trivial, since it is isomorphic to the algebra of all subclasses of a void class E. In case A is an algebra of two or more subclasses of a class E, we construct an isomorphic reduced algebra of classes B with elements which are classes of subclasses of E. We observe that the zero element 0 of A must be the void class: for if a is any element of A, the equation $a+a=0$ identifies 0 as the symmetric difference of the class a with itself. We observe also that the union of all subclasses of E belonging to A is a non-void subclass H of E. If α is any element of H, the intersection of all those classes a in A which contain α is a non-void class $\mathrm{E}(\alpha)$ contained

† Kuratowski, *Topologie*, vol. 1, Warsaw-Lwow, 1933, p. 29.

in H and containing α. We shall now show that two such classes $E(\alpha)$ and $E(\beta)$ have an element γ in common if and only if they coincide. Let a and b be arbitrary classes belonging to A and containing α and β respectively. If $E(\alpha)E(\beta)$ contains γ then ab also contains γ. Since $ab(a+b)=0$, where 0 is the void class, the symmetric difference $a+b$ cannot contain γ; furthermore it cannot contain α or β without containing $E(\alpha)$ or $E(\beta)$ and hence γ. Since α and β belong to a and to b respectively but not to $a+b$, they must belong to ab and hence to both a and b. It thus follows that $E(\alpha)=E(\beta)$. In consequence of this result, we see that a class a belonging to A has a non-void intersection with a class $E(\alpha)$ if and only if a contains $E(\alpha)$: for if β is common to a and $E(\alpha)$, then β is common to $E(\beta)$, which is contained in a, and $E(\alpha)$; and $E(\alpha)=E(\beta)$. We denote by $\mathfrak{H}$ the class of all classes $E(\alpha)$ and by $\mathfrak{H}(a)$ the class of all classes $E(\alpha)$ which are contained in the class a belonging to A. It is evident that the correspondence $a\rightarrow\mathfrak{H}(a)$ has the following properties: $\mathfrak{H}(a)=\mathfrak{H}(b)$ if and only if $a=b$; $\mathfrak{H}(a+b)=\mathfrak{H}(a)\Delta\mathfrak{H}(b)$; $\mathfrak{H}(ab)=\mathfrak{H}(a)\mathfrak{H}(b)$. Thus the class of all classes $\mathfrak{H}(a)$ is an algebra of classes B isomorphic to A. The correspondence $a\rightarrow\mathfrak{H}(a)$ may evidently be regarded as induced by the correspondence $\alpha\rightarrow E(\alpha)$ carrying H into $\mathfrak{H}$. We still have to show that B is a reduced algebra of classes. Since every α in H is contained in some class a belonging to A, it is clear that every $E(\alpha)$ in H is contained in some class $\mathfrak{H}(a)$ belonging to B. The classes $\mathfrak{H}(a)$ containing an arbitrary $E(\alpha)$ in $\mathfrak{H}$ are those for which a is a class containing α; and hence, for any β such that $E(\beta)\neq E(\alpha)$, there is some such a with the property that a does not contain $E(\beta)$, $\mathfrak{H}(a)$ does not contain $E(\beta)$. It follows that $E(\alpha)$ is the sole element common to the classes $\mathfrak{H}(a)$ containing it. Thus B is a reduced algebra of classes in accordance with Definition 10.

Obviously, if A is any reduced algebra of classes we can construct isomorphic algebras of classes which are not reduced: we have only to replace single elements of the basic class E by two or more elements and to adjoin superfluous elements, in a reversal of the process of reduction described in the proof of Theorem 54.

We shall consider also a relation of equivalence between algebras of classes, defined as follows:

DEFINITION 11. *If A and B are algebras of subclasses of classes E_A and E_B respectively and if there exists a biunivocal correspondence between E_A and E_B which induces an isomorphism $A\longleftrightarrow B$, then the algebras A and B are said to be equivalent.*

Concerning this relation of equivalence, we may state the following theorem without formal proof:

THEOREM 55. *If A, B, and C are algebras of classes, then the relation of equivalence introduced in Definition* 11 *has the following properties*:

(1) *A is equivalent to A;*

(2) *if A is equivalent to B, then B is equivalent to A;*

(3) *if A is equivalent to B and B to C, then A is equivalent to C;*

(4) *if A is a reduced algebra of classes and B is equivalent to A, then B is a reduced algebra of classes.*

It is not true that two isomorphic reduced algebras of classes are necessarily equivalent. Illustrative examples are easily constructed on the basis of the following result:

THEOREM 56. *If A is a reduced algebra of subclasses a of a class* E *and if* H *is any subclass of* E, *then the classes* Ha *constitute a reduced algebra B of subclasses of* H *homomorphic to A under the correspondence $a \to$ Ha. This homomorphism is an isomorphism if and only if the intersection* Ha *is non-void for every non-void a in A.*

The relations $\mathrm{H}(a+b) = (\mathrm{H}a)+(\mathrm{H}b)$, $\mathrm{H}(ab) = (\mathrm{H}a)(\mathrm{H}b)$ hold in the algebra of all subclasses of E and show that the correspondence $a \to \mathrm{H}a$ determines a homomorphism $A \to B$. It is evident that B is a reduced algebra of classes. If the indicated homomorphism is an isomorphism, then $\mathrm{H}a = \mathrm{H}b$ implies $a = b$; in particular, for $b = 0$, we must have $a = 0$ whenever $\mathrm{H}a = 0$. On the other hand, if $\mathrm{H}(a+b) = 0$ implies $a+b = 0$, we see that $\mathrm{H}a = \mathrm{H}b$ implies $\mathrm{H}(a+b) = \mathrm{H}a + \mathrm{H}b = 0$ and hence $a+b = 0$, or, equivalently, $a = b$; and the indicated homomorphism is therefore an isomorphism.

In view of Theorems 54 and 55, we may properly confine our attention to reduced algebras of classes; and we may regard equivalent algebras of classes as abstractly identical. We shall therefore assume in the remainder of the present chapter that we are dealing with reduced algebras of classes exclusively.

3. **The analysis of algebras of classes.** If A is an algebra of subclasses of a class E and if $\mathfrak{a}$ is an ideal in A, we can form the union $\mathrm{E}(\mathfrak{a})$ of all those subclasses of E which belong to the ideal $\mathfrak{a}$. On the other hand, if H is an arbitrary subclass of E, we can form the ideal $\mathfrak{a}(\mathrm{H})$ in A consisting of all those classes which belong to A and are contained in H. The chief task of the present section is to study the two correspondences thus defined between ideals in A and subclasses of E.

We commence with the following theorem:

THEOREM 57. *Let A be an algebra of subclasses a of a class* E; *let $\mathfrak{a}$ be an arbitrary ideal in A; and let* $\mathrm{E}(\mathfrak{a})$ *be the union of all those subclasses of* E *which are elements of the ideal $\mathfrak{a}$. Then the following relations are valid*:

(1) *if* $\mathfrak{B}$ *is any non-void class of ideals in* A, *then*

$$\sum_{\mathfrak{a}\in\mathfrak{B}} E(\mathfrak{a}) = E(S_{\mathfrak{a}\in\mathfrak{B}}\,\mathfrak{a}), \qquad \prod_{\mathfrak{a}\in\mathfrak{B}} E(\mathfrak{a}) \supset E(P_{\mathfrak{a}\in\mathfrak{B}}\,\mathfrak{a});$$

(2) *if* $\mathfrak{a}$ *and* $\mathfrak{b}$ *are ideals in* A, *then* $E(\mathfrak{ab}) = E(\mathfrak{a})E(\mathfrak{b})$;

(3) *if* $\mathfrak{a}$ *and* $\mathfrak{b}$ *are ideals in* A, *then* $\mathfrak{a} \subset \mathfrak{b}$ *implies* $E(\mathfrak{a}) \subset E(\mathfrak{b})$;

(4) *if* $\mathfrak{a}$ *and* $\mathfrak{b}$ *are ideals in* A, *then* $E(\mathfrak{a}) = E(\mathfrak{b})$ *implies* $\mathfrak{a}' = \mathfrak{b}'$;

(5) *the ideal* $\mathfrak{a}'$ *consists of those and only those subclasses of* $\mathbf{E}$ *which belong to* A *and are contained in* $E'(\mathfrak{a})$; *and* $E(\mathfrak{a}') \subset E'(\mathfrak{a})$.

The correspondence $\mathfrak{a}\to E(\mathfrak{a})$ *defines a homomorphism from the system* $\mathfrak{J}$ *of all ideals in* A (*with unrestricted addition and finite multiplication as operations*) *to the system of all classes* $E(\mathfrak{a})$ (*with the operations of forming arbitrary unions and finite intersections*), *in accordance with* (1) *and* (2) *above. This correspondence has the following special properties*:

(6) *if* $\mathfrak{a}$ *is a principal ideal* $\mathfrak{a}(a)$, *then* $E(\mathfrak{a}(a)) = a$;

(7) *if* $\mathfrak{a}$ *is a simple ideal, then* $E(\mathfrak{a}') = E'(\mathfrak{a})$;

(8) *if* $\mathfrak{a}$ *and* $\mathfrak{b}$ *are normal ideals, then* $E(\mathfrak{a}) = E(\mathfrak{b})$ *implies* $\mathfrak{a} = \mathfrak{b}$;

(9) *if* $\mathfrak{p}$ *is a prime ideal, then* $E'(\mathfrak{p})$ *contains at most one element.*

If the correspondence $\mathfrak{a}\to E(\mathfrak{a})$ *is restricted to normal ideals it is biunivocal; if it is restricted to simple, semiprincipal, or principal ideals, it defines an isomorphism and the corresponding classes* $E(\mathfrak{a})$ *constitute an algebra of classes.*

Properties (3) and (6) are evident from the definition of the class $E(\mathfrak{a})$. From (3), we have

$$\sum_{\mathfrak{a}\in\mathfrak{B}} E(\mathfrak{a}) \subset E(S_{\mathfrak{a}\in\mathfrak{B}}\,\mathfrak{a}), \qquad \prod_{\mathfrak{a}\in\mathfrak{B}} E(\mathfrak{a}) \supset E(P_{\mathfrak{a}\in\mathfrak{B}}\,\mathfrak{a}).$$

To complete the proof of (1) we must therefore show that

$$\sum_{\mathfrak{a}\in\mathfrak{B}} E(\mathfrak{a}) \supset E(S_{\mathfrak{a}\in\mathfrak{B}}\,\mathfrak{a})$$

or, equivalently, that

$$\sum_{\mathfrak{a}\in\mathfrak{B}} E(\mathfrak{a}) \supset a \text{ for every } a \text{ in } S_{\mathfrak{a}\in\mathfrak{B}}\,\mathfrak{a}.$$

Now by Theorem 17 such an element a can be expressed in the form $a = a_1 \vee \cdots \vee a_n$ where $a_1, \cdots, a_n$ belong respectively to ideals $\mathfrak{a}_1, \cdots, \mathfrak{a}_n$ in $\mathfrak{B}$; and we thus have

$$\sum_{\mathfrak{a}\in\mathfrak{B}} E(\mathfrak{a}) \supset E(\mathfrak{a}_1) \cup \cdots \cup E(\mathfrak{a}_n) \supset a_1 \vee \cdots \vee a_n = a,$$

as we wished to prove. In order to prove (2), it is now sufficient to show that $E(\mathfrak{a})E(\mathfrak{b}) \subset E(\mathfrak{ab})$. We have

$$\mathrm{E}(\mathfrak{a})\mathrm{E}(\mathfrak{b}) = \sum_{a\in\mathfrak{a}} a \sum_{b\in\mathfrak{b}} b = \sum_{\substack{a\in\mathfrak{a}\\ b\in\mathfrak{b}}} ab \subset \mathrm{E}(\mathfrak{ab})$$

since ab evidently belongs to the ideal $\mathfrak{ab}$. We next prove (5). In order that the class a be an element of the ideal $\mathfrak{a}'$ it is necessary and sufficient that $\mathfrak{a}(a)\mathfrak{a}=\mathfrak{o}$. By (2) and (6), the latter relation holds if and only if $a\mathrm{E}(\mathfrak{a})=\mathrm{E}(\mathfrak{a}(a))\mathrm{E}(\mathfrak{a})=\mathrm{E}(\mathfrak{a}(a)\mathfrak{a})=\mathrm{E}(\mathfrak{o})=0$, where 0 is the void class and the zero element of A. We thus conclude that $a\subset\mathrm{E}'(\mathfrak{a})$ and that $\mathrm{E}(\mathfrak{a}')\subset\mathrm{E}'(\mathfrak{a})$. The first part of (5) obviously implies (4). If we combine (4) with a result noted in Theorem 28, we obtain (8). To prove (7), we apply (1) and (2) to the relations $\mathfrak{a}\vee\mathfrak{a}'=\mathfrak{e}$, $\mathfrak{a}\mathfrak{a}'=\mathfrak{o}$, which hold when $\mathfrak{a}$ is simple, finding that $\mathrm{E}(\mathfrak{a})\cup\mathrm{E}(\mathfrak{a}')=\mathrm{E}(\mathfrak{a}\vee\mathfrak{a}')=\mathrm{E}(\mathfrak{e})=\mathrm{E}$, $\mathrm{E}(\mathfrak{a})\mathrm{E}(\mathfrak{a}')=\mathrm{E}(\mathfrak{a}\mathfrak{a}')=\mathrm{E}(\mathfrak{o})=0$ and hence that $\mathrm{E}(\mathfrak{a}')=\mathrm{E}'(\mathfrak{a})$. Finally, we consider (9). If $\mathrm{E}'(\mathfrak{p})$ contains elements α and β where $\alpha\neq\beta$, the fact that A is a reduced algebra of classes shows that there exists a class a in A which contains α but not β. Since a does not belong to $\mathfrak{p}$, we have $\mathfrak{p}\vee\mathfrak{a}(a)=\mathfrak{e}$ in accordance with Theorem 39. Thus we see that $\mathrm{E}(\mathfrak{p})\cup a=\mathrm{E}(\mathfrak{p})\cup\mathrm{E}(\mathfrak{a}(a))=\mathrm{E}(\mathfrak{p}\vee\mathfrak{a}(a))=\mathrm{E}(\mathfrak{e})=\mathrm{E}$. Since β is not an element of $\mathrm{E}(\mathfrak{p})$, it must be an element of a, contrary to hypothesis. Hence (9) is established. It is now evident that the correspondence $\mathfrak{a}\to\mathrm{E}(\mathfrak{a})$ sets up a homomorphism as stated in the theorem. We note that the congruence determined in $\mathfrak{Z}$ by this homomorphism, namely, the relation $\mathfrak{a}\equiv\mathfrak{b}$ equivalent to $\mathrm{E}(\mathfrak{a})=\mathrm{E}(\mathfrak{b})$, has by (4) the property that it implies the congruence C of Theorem 38. It follows that the system of classes $\mathrm{E}(\mathfrak{a})$ has the system $\mathfrak{Z}^C$ of Theorem 28 as a homomorph. The specialization of this correspondence to the various special classes of ideals now follows at once from Theorems 28–32 and Theorem 42.

THEOREM 58. *Let A be an algebra of subclasses a of a class* E; *let* H *be an arbitrary subclass of* E; *and let* $\mathfrak{a}$(H) *be the class of all elements a in A which are subclasses of* H. *Then* $\mathfrak{a}$(H) *is an ideal in A with the following properties*:

(1) $\mathfrak{a}(\mathrm{H}_1)\vee\mathfrak{a}(\mathrm{H}_2)\subset\mathfrak{a}(\mathrm{H}_1\cup\mathrm{H}_2)$;

(2) $\mathfrak{a}(\mathrm{H}_1)\mathfrak{a}(\mathrm{H}_2)=\mathfrak{a}(\mathrm{H}_1\mathrm{H}_2)$;

(3) $\mathrm{H}_1\subset\mathrm{H}_2$ *implies* $\mathfrak{a}(\mathrm{H}_1)\subset\mathfrak{a}(\mathrm{H}_2)$;

(4) $\mathfrak{a}$(H) *is prime if and only if* H′ *has exactly one element.*

In connection with Theorem 57, *the following relations are found to hold*:

(5) $\mathfrak{a}(\mathrm{E}(\mathfrak{b}))\supset\mathfrak{b}$; (6) $\mathrm{E}(\mathfrak{a}(\mathrm{H}))\subset\mathrm{H}$; (7) $\mathfrak{b}'=\mathfrak{a}(\mathrm{E}'(\mathfrak{b}))$.

It is easily verified that $\mathfrak{a}$(H), which always contains the void class 0, that is, the zero element in A, is an ideal in A. Properties (1), (3), (5), and (6) are evident; and (7) is a restatement of Theorem 57 (5). Since the relation

$\mathfrak{a}(H_1)\mathfrak{a}(H_2) \supset \mathfrak{a}(H_1H_2)$ follows from (3), we can prove (2) by showing that $\mathfrak{a}(H_1)\mathfrak{a}(H_2)$ is contained in $\mathfrak{a}(H_1H_2)$. If a is in the ideal $\mathfrak{a}(H_1)\mathfrak{a}(H_2)$, Theorem 18 shows that $a = bc$, where b and c are in $\mathfrak{a}(H_1)$ and $\mathfrak{a}(H_2)$ respectively. It is thus clear that a is contained in H_1H_2 and hence that $\mathfrak{a}(H_1)\mathfrak{a}(H_2) \subset \mathfrak{a}(H_1H_2)$, as we wished to prove. Finally, we consider (4). If $\mathfrak{a}(H)$ is prime, then $E'(\mathfrak{a}(H))$ contains at most one element by Theorem 57 (9). Hence H' contains at most one element by (6) above. It is clear, however, that H' must contain at least one element, since $H' = 0$ would imply $\mathfrak{a}(H) = \mathfrak{a}(E) = \mathfrak{e}$, contrary to hypothesis. On the other hand, if H' contains exactly one element α, $\mathfrak{a}(H)$ consists of those classes in A which do not contain α; and the remaining classes in A all contain α and constitute a class $\mathfrak{b}$, which cannot be void since A is a reduced algebra of classes. It is evident that the classes $\mathfrak{a}(H)$ and $\mathfrak{b}$ have the properties enumerated in Theorem 34, and hence that $\mathfrak{a}(H)$ is a prime ideal.

It is now desirable that we consider in greater detail two special cases corresponding to the possible extremes under Theorem 57 (4): the cases where the relation $E(\mathfrak{a}) = E(\mathfrak{b})$ is equivalent to the relations $\mathfrak{a} = \mathfrak{b}$, and $\mathfrak{a}' = \mathfrak{b}'$ respectively. For subsequent considerations, it is helpful to introduce the following terminology in the first case:

DEFINITION 12. *An algebra A of subclasses of a class* E *is said to be perfect if* $E(\mathfrak{a}) = E(\mathfrak{b})$ *implies* $\mathfrak{a} = \mathfrak{b}$; *that is, if the homomorphism of Theorem* 57 *is an isomorphism.*

The characterization of perfect algebras of classes involves the following connection with the closing remarks of Chapter II, §4:

THEOREM 59. *In order that an algebra A of subclasses of a class* E *be perfect, it is necessary and sufficient that*

(1) *the Fundamental Proposition of Ideal Arithmetic hold in A;*

(2) $E'(\mathfrak{p})$ *be a one-element class whenever* $\mathfrak{p}$ *is a prime ideal.*

If the algebra A is perfect then $E(\mathfrak{a}) = E = E(\mathfrak{e})$ implies $\mathfrak{a} = \mathfrak{e}$. Theorem 57 (9) therefore shows, by virtue of the relation $E(\mathfrak{p}) \neq \mathfrak{e}$, that $E'(\mathfrak{p})$ is a one-element class whenever $\mathfrak{p}$ is a prime ideal. If $\mathfrak{a}$ is an arbitrary ideal not equal to $\mathfrak{e}$, then the prime ideal $\mathfrak{p}$ is a divisor of $\mathfrak{a}$ if and only if $\mathfrak{p}\mathfrak{a} = \mathfrak{a}$; the latter relation is equivalent to $E(\mathfrak{p}\mathfrak{a}) = E(\mathfrak{a})$, since A is perfect, and hence to the relations $E(\mathfrak{p})E(\mathfrak{a}) = E(\mathfrak{p}\mathfrak{a}) = E(\mathfrak{a})$, $E(\mathfrak{a}) \subset E(\mathfrak{p})$. If α is an element not contained in $E(\mathfrak{a}) \neq E$, the class H of all elements in E not equal to α determines a prime ideal $\mathfrak{p} = \mathfrak{a}(H)$ in accordance with Theorem 58 (4). By Theorem 58 (6), $E(\mathfrak{p}) \subset H$; since $E'(\mathfrak{p})$ is a one-element class by the preceding results, it follows that $E(\mathfrak{p}) = H \supset E(\mathfrak{a})$ and hence that $\mathfrak{p} \supset \mathfrak{a}$. Thus the class $\mathfrak{B}$ of all prime ideal divisors of $\mathfrak{a}$ is non-void and consists of the ideals $\mathfrak{p} = \mathfrak{a}(H)$ where $H \supset E(\mathfrak{a})$ and

H' is a one-element class. We can now show that $\mathfrak{a} = P_{\mathfrak{p} \in \mathfrak{B}} \mathfrak{p}$. By Theorem 57 (1) and the relation $E(\mathfrak{p}) = H$, we have $E(P_{\mathfrak{p} \in \mathfrak{B}} \mathfrak{p}) \subset \Pi_{\mathfrak{p} \in \mathfrak{B}} E(\mathfrak{p}) = E(\mathfrak{a})$. On the other hand, the relation $\mathfrak{a} \subset P_{\mathfrak{p} \in \mathfrak{B}} \mathfrak{p}$ shows that $E(\mathfrak{a}) \subset E(P_{\mathfrak{p} \in \mathfrak{B}} \mathfrak{p})$. It follows that $E(\mathfrak{a}) = E(P_{\mathfrak{p} \in \mathfrak{B}} \mathfrak{p})$ and, since A is perfect, that $\mathfrak{a} = P_{\mathfrak{p} \in \mathfrak{B}} \mathfrak{p}$, as we wished to prove.

Now let us suppose that A is an algebra of classes in which conditions (1) and (2) are satisfied. If $\mathfrak{a}$ and $\mathfrak{b}$ are distinct ideals, there must exist a prime ideal $\mathfrak{p}$ containing one but not both of these ideals, in accordance with (1). Let us suppose that our notation is so chosen that $\mathfrak{p}$ contains $\mathfrak{a}$ but not $\mathfrak{b}$. By Theorem 57 (3) we have $E(\mathfrak{p}) \supset E(\mathfrak{a})$. On the other hand, (2) shows that $E'(\mathfrak{p})$ is a one-element class and hence that $\mathfrak{a}(E(\mathfrak{p}))$ is a prime ideal divisor of $\mathfrak{p}$ in accordance with Theorem 58, (4) and (5). It follows that $\mathfrak{p} = \mathfrak{a}(E(\mathfrak{p}))$. If the relation $E(\mathfrak{p}) \supset E(\mathfrak{b})$ were true, we should have $\mathfrak{b} \subset \mathfrak{a}(E(\mathfrak{b})) \subset \mathfrak{a}(E(\mathfrak{p})) = \mathfrak{p}$, contrary to hypothesis. Thus $E(\mathfrak{p})$ contains $E(\mathfrak{a})$ but not $E(\mathfrak{b})$, so that $E(\mathfrak{a}) \neq E(\mathfrak{b})$. We therefore conclude that A is perfect.

We now turn to the second case, that in which $E(\mathfrak{a}) = E(\mathfrak{b})$ and $\mathfrak{a}' = \mathfrak{b}'$ are equivalent relations. In view of Theorem 57 (4), we have merely to ascertain the conditions under which $\mathfrak{a}' = \mathfrak{b}'$ implies $E(\mathfrak{a}) = E(\mathfrak{b})$.

THEOREM 60. *The following assertions concerning an algebra A of subclasses of a class* E *are equivalent*:

(1) $\mathfrak{a}' = \mathfrak{b}'$ *implies* $E(\mathfrak{a}) = E(\mathfrak{b})$;

(2) *every one-element subclass of* E *is an element of* A;

(3) $\mathfrak{a}(H_1) = \mathfrak{a}(H_2)$ *implies* $H_1 = H_2$;

(4) $E(\mathfrak{a}(H)) = H$ *for every subclass* H *of* E;

(5) $E(\mathfrak{a}') = E'(\mathfrak{a})$ *for every ideal* $\mathfrak{a}$ *in* A;

(6) $\mathfrak{a}(H)$ *is a normal ideal for every* H;

(7) $\mathfrak{a}(E(\mathfrak{b})) = \mathfrak{b}''$ *for every ideal* $\mathfrak{b}$ *in* A;

(8) $\mathfrak{a}(H_1) \vee \mathfrak{a}(H_2) = \mathfrak{a}(H_1 \cup H_2)$ *for all* H_1 *and* H_2.

When these conditions are satisfied, the Boolean ring $\mathfrak{N}$ *of all normal ideals in* A *described in Theorem* 29 *is isomorphic to the algebra of all subclasses of* E. *Conversely, if the Boolean ring* $\mathfrak{N}$ *of all normal ideals in an abstract Boolean ring* B *is isomorphic to the algebra of all subclasses of a class* E, *then* B *is isomorphic to a reduced algebra* A *of subclasses of* E *in which these conditions are satisfied.*

We begin by establishing the implications $1 \to 2 \to 3 \to 4 \to 5 \to 6 \to 7 \to 1$, $4 \to 8$, and $8 \to 6$. If (1) holds and H is a one-element subclass of E, then $\mathfrak{a}(H')$ is prime by Theorem 58 (4), $E(\mathfrak{a}(H')) = H'$ by Theorem 57 (9) and Theorem 58 (6), and $\mathfrak{a}'(H') \neq \mathfrak{o}$ since (1) and $\mathfrak{a}'(H') = \mathfrak{e}'$ would imply $H' = E(\mathfrak{a}(H')) = E(\mathfrak{e}) = E$; and it is therefore evident that H, as the only subclass of H distinct from 0, must be an element of A. If (2) holds, the ideal $\mathfrak{a}(H)$ contains those and only those one-element classes which are subclasses of H; it is thus clear that

$\mathfrak{a}(H_1)=\mathfrak{a}(H_2)$ implies $H_1=H_2$. If (3) holds, we have $\mathfrak{a}(E(\mathfrak{a}(H)))=\mathfrak{a}(H)$ by Theorem 58, (3), (5), and (6), and hence $E(\mathfrak{a}(H))=H$ for every H. If (4) holds, we have $\mathfrak{a}'=\mathfrak{a}(E'(\mathfrak{a}))$ by Theorem 58 (7), and hence $E(\mathfrak{a}')=E(\mathfrak{a}(E'(\mathfrak{a})))$, so that $E(\mathfrak{a}')=E'(\mathfrak{a})$. If (5) holds, we have $E(\mathfrak{a}''(H))=E'(\mathfrak{a}'(H))=E''(\mathfrak{a}(H))=E(\mathfrak{a}(H))$ and hence $\mathfrak{a}''(H)\subset\mathfrak{a}(E(\mathfrak{a}''(H)))=\mathfrak{a}(E(\mathfrak{a}(H)))=\mathfrak{a}(H)\subset\mathfrak{a}''(H)$ by Theorem 58, (3), (5), and (6), so that $\mathfrak{a}(H)=\mathfrak{a}''(H)$ and $\mathfrak{a}(H)$ is normal for every H. If (6) holds, we have $\mathfrak{a}(E(\mathfrak{b}))\mathfrak{b}'=\mathfrak{a}(E(\mathfrak{b}))\mathfrak{a}(E'(\mathfrak{b}))=\mathfrak{a}(E(\mathfrak{b})E'(\mathfrak{b}))=\mathfrak{a}(0)=\mathfrak{o}$, $\mathfrak{a}(E(\mathfrak{b}))\supset\mathfrak{b}$ by Theorem 58, (2), (5), and (7), and hence $\mathfrak{b}''\supset\mathfrak{a}(E(\mathfrak{b}))=\mathfrak{a}''(E(\mathfrak{b}))\supset\mathfrak{b}''$, so that $\mathfrak{a}(E(\mathfrak{b}))=\mathfrak{b}''$. If (7) holds, then $E(\mathfrak{b})=E(\mathfrak{a}(E(\mathfrak{b})))=E(\mathfrak{b}'')$ by Theorem 57 (3) and Theorem 58, (5) and (6), so that $\mathfrak{a}'=\mathfrak{b}'$ implies $\mathfrak{a}''=\mathfrak{b}''$ and $E(\mathfrak{a})=E(\mathfrak{a}'')=E(\mathfrak{b}'')=E(\mathfrak{b})$. If (4) holds, we have $E(\mathfrak{a}(H_1\cup H_2))=H_1\cup H_2=E(\mathfrak{a}(H_1))\cup E(\mathfrak{a}(H_2))$ by Theorem 57 (1) and hence $\mathfrak{a}'(H_1\cup H_2)=(\mathfrak{a}(H_1)\vee\mathfrak{a}(H_2))'$ by Theorem 57 (4); and, since (4) implies (6), we have $\mathfrak{a}(H_1)\nabla\mathfrak{a}(H_2)=(\mathfrak{a}(H_1)\vee\mathfrak{a}(H_2))''=\mathfrak{a}''(H_1\cup H_2)=\mathfrak{a}(H_1\cup H_2)$, in accordance with Definition 9. If (8) holds, we have $\mathfrak{a}''(H)=\mathfrak{a}(H)\nabla\mathfrak{a}(0)=\mathfrak{a}(H\cup 0)=\mathfrak{a}(H)$. The equivalence of assertions (1)–(8) is thus established.

The correspondence $H\rightarrow\mathfrak{a}(H)$ takes the class of all subclasses of E into a subclass of the Boolean ring $\mathfrak{R}$ of all normal ideals in A by virtue of (6); and does so in a biunivocal manner in accordance with (3). If $\mathfrak{b}$ is any normal ideal, and $H=E(\mathfrak{b})$, then $\mathfrak{a}(H)=\mathfrak{a}(E(\mathfrak{b}))=\mathfrak{b}''=\mathfrak{b}$ by (7), so that the correspondence takes the class of all subclasses of E into the entire class $\mathfrak{R}$. By Theorem 58 (2) and (8) above, this correspondence determines the isomorphism described in the statement of the theorem. Let us suppose conversely that the Boolean ring $\mathfrak{R}$ associated with an abstract Boolean ring B is isomorphic to the algebra of all subclasses of a class E. Since B is isomorphic to the subring of principal ideals $\mathfrak{P}$, it is evident that B is also isomorphic to an algebra A of subclasses of E. If H is an arbitrary one-element subclass of E, it corresponds in the isomorphism to a normal ideal $\mathfrak{a}$ distinct from 0. Since every element of B which belongs to the ideal $\mathfrak{a}$ corresponds to a subclass of H, we see that $\mathfrak{a}$ must consist of exactly two elements corresponding respectively to H and to the void class. It follows that H is an element of the algebra A. Since A has property (2) of the present theorem by the result just proved, A is evidently a reduced algebra of classes in accordance with Definition 10; and it has all the other properties (1)–(8).

We see that Theorem 60 can be made to yield information about abstract Boolean rings by virtue of the following result:

THEOREM 61. *In order that an abstract Boolean ring B be isomorphic to an algebra A of subclasses of a class* E *with the property* (2) *of Theorem* 60, *it is necessary and sufficient that B contain a complete atomic system.*

The sufficiency of the condition is evident from Theorem 8. The necessity follows at once from the observation that the one-element subclasses of E constitute a complete atomic system in A in accordance with Definition 5: for the image of this system in B under the isomorphism $A \longleftrightarrow B$ is also a complete atomic system.

We have on the basis of Theorems 60 and 61 the following proposition:

THEOREM 62. *In order that an abstract Boolean ring B be isomorphic to the algebra A of all subclasses of a class* E *it is necessary and sufficient that every normal ideal in B be principal and that B contain a complete atomic system.*

The existence of a complete atomic system in B is necessary and sufficient for the existence of an isomorphism $B \longleftrightarrow A$, where A is an algebra of classes with property (2) of Theorem 60, as we have just proved in Theorem 61. In order that every subclass of the basic class belong to such an isomorphic algebra A, it is necessary and sufficient, by virtue of Theorem 60, that every normal ideal in A be principal: for the normal ideals in A are precisely the ideals $\mathfrak{a}(H)$; and such an ideal is principal if and only if the generating class H is an element of A. The isomorphism $A \longleftrightarrow B$ shows that A has the indicated property if and only if every normal ideal in B is principal. The proof is thus completed.

We may remark that Theorems 60 and 61 furnish the basis for our comments concerning the class $\mathfrak{N}$ of normal ideals at the end of Chapter II, §3. Since a Boolean ring does not necessarily contain any atomic element, let alone a complete atomic system, $\mathfrak{N}$ cannot in general be isomorphic to an algebra of all subclasses of an appropriate fixed class.

4. **An illustration.** We shall now consider, by way of illustration, the application of the results of §3 to the algebra A of all Lebesgue measurable subsets of finite measure in the euclidean plane E. Since every one-element set belongs to A, we see that A has property (2) of Theorem 60. Hence the normal ideals in A are precisely the ideals $\mathfrak{a}(H)$, consisting of all sets belonging to A and contained in the arbitrary set H. In order that an ideal be principal, it must be normal and hence must be expressible in the form $\mathfrak{a}(H)$ where H is an element of A. We can now determine the simple ideals in A on the basis of Theorem 26 as follows: every simple ideal is normal and can therefore be expressed as $\mathfrak{a}(H)$; in order that $\mathfrak{a}(H)$ be simple, the product $\mathfrak{a}(H)\mathfrak{a}(H_1) = \mathfrak{a}(HH_1)$ must be principal whenever $\mathfrak{a}(H_1)$ is principal, and conversely; in other words $\mathfrak{a}(H)$ is simple if and only if H has the characteristic property of Lebesgue measurable sets, that HH_1 is in A whenever H_1 is in A. The semiprincipal ideals which are not principal are now seen to be the ideals $\mathfrak{a}(H)$ where the complement of H belongs to A. From these characterizations of the ideals in A, we

see that the classes $\mathfrak{P}$, $\mathfrak{P}^*$, $\mathfrak{S}$, and $\mathfrak{N}$ are distinct. Furthermore, since the sets of zero measure constitute an ideal in A which obviously cannot be put in the form $\mathfrak{a}(H)$, the classes $\mathfrak{N}$ and $\mathfrak{Z}$ are also distinct. The prime ideals belonging to $\mathfrak{N}$ are evidently given by $\mathfrak{a}(H)$ where the complement of H is a one-element class. Hence no normal prime ideal in A can contain every set of zero measure; and, if the ideal of sets of zero measure has a prime ideal divisor in A, the latter ideal cannot be normal. This familiar example therefore shows that the classifications discussed in Chapter II, §§3 and 4, are not degenerate except in special Boolean rings.

Chapter IV. Representation theory

1. **General remarks.** We come now to the problem of constructing an algebra of classes isomorphic to a given abstract Boolean ring. Indeed, we shall consider the broader problem of determining all algebras of classes homomorphic to a given Boolean ring. In view of the results of Chapter III, §2, we may confine our attention to reduced algebras of classes and we may regard representations by equivalent algebras of classes as abstractly identical. Thus the general problem of representation theory can be formulated in the following precise terms: if A is any Boolean ring, it is required to construct a family of reduced algebras of classes homomorphic to A so that any algebra of classes homomorphic to A is equivalent, after the reduction described in Theorem 54, to an algebra belonging to this family. We shall be able to solve this problem completely in the present chapter.

The results of Chapter III, §3, are sufficient to indicate that the representation problem is closely bound up with the theory of ideals, particularly the prime ideals. Suppose that the Boolean ring A has a homomorph B with elements which are subclasses of a fixed class E. Then each element of E is associated by Theorem 58 (4) with a prime ideal in B and hence, by virtue of Theorem 48, with a prime ideal in A. We may therefore suppose that each element of E is replaced by the associated prime ideal, so that A is represented by an algebra of classes of its own prime ideals. If this representation be further analyzed, it becomes clear that a given element a in A must be represented by the class of those prime ideals which are thus determined by the given basic class E and do not contain a as an element. Thus any attempt to solve the representation problem must be based upon a preliminary theory of the existence and divisibility properties of prime ideals. In particular, in order to show that A has an isomorphic representation, it is evidently necessary to prove that whenever a and b are unequal elements of A there exists a prime ideal containing one, but not both, of these elements. It is not difficult to see that the existence theorem thus suggested is equivalent to the Funda-

mental Proposition of Ideal Arithmetic, stated at the close of Chapter II, §4.

Our program for solving the representation problem is thus prescribed: we first prove the Fundamental Existence Proposition and the Fundamental Proposition of Ideal Arithmetic, as previously stated; we then determine a representation by subclasses of the class $\mathfrak{S}$ of all prime ideals, noting that this algebra of classes must be perfect in accordance with Definition 12 and Theorem 59; and we then show that all other representations can be constructed from this perfect representation by the process discussed in Theorem 56. It will be observed that the indicated perfect representation also affords an isomorphic representation of the class of all ideals under the operations of unrestricted addition and finite multiplication, in accordance with Theorem 57 and Definition 12.†

It will be evident from a closer inspection of our solution of the representation problem, that the relations thereby revealed are of such a nature that they can appropriately be stated in topological terms. Once such a statement has been made, a number of interesting topological connections are suggested. The further development of the theory in the indicated direction will be carried out in another paper.‡

2. **Existence and divisibility properties of prime ideals.** In view of the remarks made in the preceding section, we shall now resume the study of prime ideals at the point where we broke off in Chapter II, §4. We first establish the Fundamental Existence Proposition.

THEOREM 63. *In a Boolean ring A containing at least two elements, there exists at least one prime ideal.*

We shall give two proofs, both based upon the principle of transfinite induction.§ In view of Theorems 34–36 and 39, we see that the construction of a prime ideal, or, alternatively, the construction of a homomorphism $A \rightarrow B$

† The construction of this perfect representation was described in a communication to the Society, abstracted in its Bulletin, abstract 39-3-86, received January 24, 1933, and also in Proceedings of the National Academy of Sciences, vol. 20 (1934), pp. 197–202. A more general representation theorem (for C-lattices or distributive lattices) was discovered independently and only slightly later by Garrett Birkhoff, see Proceedings of the Cambridge Philosophical Society, vol. 29 (1933), pp. 441–464, Theorem 25.2. Since MacNeille has shown in his doctoral dissertation *The Theory of Partially Ordered Sets* (1935), not yet published, that every distributive lattice can be algebraically imbedded in a Boolean algebra, or Boolean ring, Birkhoff's result may now be regarded as included in those given here.

‡ See Stone, Proceedings of the National Academy of Sciences, vol. 20 (1934), pp. 197–202.

§ Still another proof is given by von Neumann and Stone, Fundamenta Mathematicae, vol. 25 (1935), pp. 353–378; a paper by Tarski, Fundamenta Mathematicae, vol. 15 (1930), pp. 42–50, although couched in the language of the theory of measure, also gives a proof which could be adapted to the present situation.

where B is a two-element Boolean ring necessitates the distribution of the elements of A between two subclasses according to certain prescribed rules, and hence requires the examination and allocation of those elements to subclasses in some orderly fashion. In general, therefore, we can hardly expect to devise a method of construction which does not involve the well-ordering hypothesis.

In our first proof, we work with the class $\mathfrak{J}$ of all ideals in A. We choose an ordinal number ω so that the ideals in A can be put in biunivocal correspondence with the ordinals γ such that $\gamma<\omega$; and we may even suppose that ω is the first ordinal number available for this purpose. Since A has more than one element, we can select an element a not equal to 0. We now define a prime ideal $\mathfrak{p}$ by transfinite induction, first forming ideals $\mathfrak{b}_\alpha$ for all $\alpha<\omega$ and then putting $\mathfrak{p}=S_{\alpha<\omega}\mathfrak{b}_\alpha$. If $\mathfrak{a}_1$ is the first ideal in $\mathfrak{J}$, we define $\mathfrak{b}_1$ as equal to $\mathfrak{o}$ or to $\mathfrak{a}_1$, according as a belongs to $\mathfrak{a}_1$ or not; and, if $\mathfrak{b}_\alpha$ has been defined for all ordinals α such that $\alpha<\beta$ where $\beta<\omega$, we define $\mathfrak{b}_\beta$ as the ideal $S_{\alpha<\beta}\mathfrak{b}_\alpha$ or as the ideal $\mathfrak{a}_\beta \vee S_{\alpha<\beta}\mathfrak{b}_\alpha$ according as a belongs to the latter ideal or not. By the principle of transfinite induction, $\mathfrak{p}$ exists and is an ideal in A. We shall now show that $\mathfrak{p}$ is divisorless and hence prime in accordance with Theorem 33. We first prove that $\mathfrak{p}\neq\mathfrak{e}$; in particular, that $\mathfrak{p}$ does not contain the element a. Since $\alpha<\beta$ implies $\mathfrak{b}_\alpha \subset \mathfrak{b}_\beta$ by construction, we see that, if $\mathfrak{p}$ contained a, Theorem 17 would establish the existence of ordinals $\beta_1<\cdots<\beta_n<\omega$ such that a belongs to $\mathfrak{b}_{\beta_n}=\mathfrak{b}_{\beta_1}\vee\cdots\vee\mathfrak{b}_{\beta_n}$. Similarly, if $\mathfrak{b}_\beta$ contained the element a for $\beta>1$, Theorem 17 would establish the existence of an ordinal $\alpha<\beta$ such that $\mathfrak{b}_\alpha$ contains a: for the fact that $\mathfrak{b}_\beta$ is assumed to contain a excludes the relation $\mathfrak{b}_\beta=\mathfrak{a}_\beta\vee S_{\alpha<\beta}\mathfrak{b}_\alpha$ and thus implies $\mathfrak{b}_\beta=S_{\alpha<\beta}\mathfrak{b}_\alpha$. We see therefore that, if $\mathfrak{p}$ contained a, there would be a first ordinal α such that $\mathfrak{b}_\alpha$ contains a and this ordinal would necessarily be $\alpha=1$. By construction, however, $\mathfrak{b}_1$ does not contain a. Next we prove that if $\mathfrak{a}$ is an ideal divisor of $\mathfrak{p}$, then $\mathfrak{a}=\mathfrak{p}$ or $\mathfrak{a}=\mathfrak{e}$. In the first place let us suppose that $\mathfrak{a}$ does not contain a. If γ is the ordinal assigned to $\mathfrak{a}$, we then have $\mathfrak{a}=\mathfrak{a}_\gamma\supset\mathfrak{p}\supset S_{\alpha<\gamma}\mathfrak{b}_\alpha$ and hence $\mathfrak{p}\supset\mathfrak{b}_\gamma=\mathfrak{a}_\gamma\vee S_{\alpha<\gamma}\mathfrak{b}_\alpha=\mathfrak{a}_\gamma$, since $\mathfrak{a}_\gamma$ does not contain a. Thus in this case we have $\mathfrak{a}=\mathfrak{a}_\gamma=\mathfrak{p}$. In the second place we suppose that $\mathfrak{a}$ contains a and hence that $\mathfrak{a}\supset\mathfrak{p}\vee\mathfrak{a}(a)$. Since the ideal $\mathfrak{a}'(a)$ is equal to $\mathfrak{a}_\delta$ for an appropriate ordinal number δ, and since $\mathfrak{a}(a)\subset\mathfrak{a}_\delta\vee S_{\alpha<\delta}\mathfrak{b}_\alpha$ would imply

$$\mathfrak{a}(a)=\mathfrak{a}(a)\mathfrak{a}_\delta\vee\mathfrak{a}(a)\sum_{\alpha<\delta}\mathfrak{b}_\alpha=\mathfrak{a}(a)\sum_{\alpha<\delta}\mathfrak{b}_\alpha\subset\sum_{\alpha<\delta}\mathfrak{b}_\alpha\subset\mathfrak{p},$$

contrary to the fact that $\mathfrak{p}$ does not contain a, we have

$$\mathfrak{a}'(a)=\mathfrak{a}_\delta\subset\mathfrak{a}_\delta\vee\sum_{\alpha<\delta}\mathfrak{b}_\alpha=\mathfrak{b}_\delta\subset\mathfrak{p}$$

and hence

$$\mathfrak{a} \supset \mathfrak{p} \vee \mathfrak{a}(a) \supset \mathfrak{a}'(a) \vee \mathfrak{a}(a) = \mathfrak{e}, \qquad \mathfrak{a} = \mathfrak{e}.$$

The proof that $\mathfrak{p}$ is a prime ideal is thus completed. It is of interest to indicate a slight modification of the foregoing construction. The hypothesis that A has two or more elements is equivalent to the hypothesis that A contains two or more ideals. If, then, $\mathfrak{a}$ and $\mathfrak{b}$ are two ideals such that $\mathfrak{b}$ is not divisible by $\mathfrak{a}$, we may suppose that the numbering of $\mathfrak{Z}$ is arranged so that $\mathfrak{a}_1 = \mathfrak{a}$ and we may choose a as an element which is in $\mathfrak{b}$ but not in $\mathfrak{a}$ and hence is different from 0. The construction given above then yields a prime ideal $\mathfrak{p}$ which is a divisor of $\mathfrak{a}$ but not of $\mathfrak{b}$. We note also that, when A has a unit e, the proof that $\mathfrak{p}$ is a prime ideal given above assumes a slightly simpler form if one takes $a = e$ at the outset.

In our second proof, which is reminiscent of the discussion of simple extensions of a commutative field, we begin by considering the case of a Boolean ring A with unit e. Let the elements of A be placed in biunivocal correspondence with the ordinal numbers γ less than an appropriately chosen ordinal number ω, the element a corresponding to γ being denoted by a_γ. We may suppose that in this correspondence $a_1 = e$, and that a_2 is different from 0 and e when A has more than two elements and is chosen at pleasure as any element a subject to this restriction. We denote by $\mathfrak{a}_\beta$ the subring of A generated by the class of all elements a_α, $\alpha \leqq \beta$, $1 \leqq \beta < \omega$; and by $\mathfrak{b}_\beta$ the subring of A generated by the class of all elements a_α, $\alpha < \beta$, $1 < \beta < \omega$. It is then evident that $\mathfrak{b}_\alpha \subset \mathfrak{a}_\alpha \subset \mathfrak{b}_\beta \subset \mathfrak{a}_\beta$ for $\alpha < \beta$; that $\mathfrak{a}_\beta = \mathfrak{b}_\beta \vee \{a_\beta\}$, $\beta > 1$, where $\{a_\beta\}$ is the subring generated by the element a_β; and that $\mathfrak{b}_\beta = S_{\alpha<\beta}\mathfrak{a}_\alpha$, $\beta > 1$. We shall now show with the help of the principle of transfinite induction that there exists a homomorphism $A \to B$ where B is a two-element Boolean ring. The subring $\mathfrak{a}_1 = \mathfrak{b}_2$ generated by $a_1 = e \neq 0$ consists of the two elements 0 and e so that we can define a homomorphism (indeed, an isomorphism) from $\mathfrak{a}_1$ to B in just one way. Let us suppose that homomorphisms $\mathfrak{a}_\alpha \to B$ have been defined for all ordinals $\alpha < \beta$ in such a way that the correspondent in B of an element in $\mathfrak{a}_\alpha$ is independent of α. Since $\mathfrak{b}_\beta = S_{\alpha<\beta}\mathfrak{a}_\alpha$, it is then evident that the homomorphisms $\mathfrak{a}_\alpha \to B$ together determine a homomorphism $\mathfrak{b}_\beta \to B$: for, if b and c are elements of $\mathfrak{b}_\beta$, there exists an ordinal $\alpha < \beta$ such that b and c are elements of $\mathfrak{a}_\alpha$, as we see by reference to the properties of the subrings $\mathfrak{a}_\alpha$ noted above; and the homomorphism $\mathfrak{a}_\alpha \to B$ implies that the elements $b+c$, bc have the correspondents b^*+c^*, b^*c^* respectively, where b^* and c^* are the fixed correspondents in B of b and c respectively. Now the subring $\mathfrak{b}_\beta$ evidently contains the element $a_1 = e$. Theorem 14 therefore shows that the subring $\alpha_\beta = \mathfrak{b}_\beta \vee \{a_\beta\}$ consists of those and only those elements of A which are expressible in the

form $xa_\beta+ya_\beta' = xa_\beta \vee ya_\beta'$ where x and y are elements of $\mathfrak{b}_\beta$. Since we have $(x_1a_\beta+y_1a_\beta')+(x_2a_\beta+y_2a_\beta') = (x_1+x_2)a_\beta+(y_1+y_2)a_\beta'$, $(x_1a_\beta+y_1a_\beta')(x_2a_\beta+y_2a_\beta') = (x_1x_2)a_\beta+(y_1y_2)a_\beta'$, we see that $\mathfrak{a}_\beta$ is a homomorph of the Boolean ring $\mathfrak{B}_\beta$ obtained as the direct sum of $\mathfrak{b}_\beta$ with itself, this homomorphism being given explicitly by the correspondence $(x, y)\rightarrow xa_\beta+ya_\beta'$. The ideal $\mathfrak{c}_\beta$ determined by this correspondence consists of those and only those elements (x, y) in $\mathfrak{B}_\beta$ such that $xa_\beta+ya_\beta' = 0$, or, equivalently, $xa_\beta = ya_\beta' = 0$; and $\mathfrak{a}_\beta$ is isomorphic to the quotient ring $\mathfrak{B}_\beta/\mathfrak{c}_\beta$ in accordance with Theorem 43. We note that, if (x, y) is in $\mathfrak{c}_\beta$, then $xy=0$ since $xy = xye = xy(a_\beta \vee a_\beta') = xya_\beta \vee xya_\beta' = 0 \vee 0 = 0$. Now the homomorphism $\mathfrak{b}_\beta\rightarrow B$ evidently induces a homomorphism $\mathfrak{B}_\beta\rightarrow C$, where C is a four-element Boolean ring obtained as the direct sum of B with itself. The image of the ideal $\mathfrak{c}_\beta$ under the latter homomorphism is an ideal $\mathfrak{c}$ in C, by virtue of Theorem 46. The ideal $\mathfrak{c}$ cannot contain the unit element (e^*, e^*) in C since, if an element (x, y) in $\mathfrak{c}_\beta$ has the correspondent (x^*, y^*) in C, then $xy=0$ implies $x^*y^*=0^*$ (the unit and zero elements in B being denoted by e^* and 0^* respectively). The four elements of C are $(0^*, 0^*)$, $(0^*, e^*)$, $(e^*, 0^*)$, (e^*, e^*); and the ideal $\mathfrak{c}$ must clearly be a principal ideal generated by one of the three elements $(0^*, 0^*)$, $(0^*, e^*)$, $(e^*, 0^*)$. It is easily verified that the quotient ring $C/\mathfrak{c}$ is homomorphic to $\mathfrak{B}_\beta/\mathfrak{c}_\beta$ and hence also to $\mathfrak{a}_\beta$. We now have three cases to consider, according to the nature of the ideal $\mathfrak{c}$. First, if $\mathfrak{c}$ is the principal ideal generated by the element $(e^*, 0^*)$, $C/\mathfrak{c}$ is a two-element ring which is isomorphic to B in one and only one way. Now, if x is an element of $\mathfrak{b}_\beta$ with correspondent x^* in B under the homomorphism $\mathfrak{b}_\beta\rightarrow B$, the homomorphism $\mathfrak{B}_\beta\rightarrow C$ takes (x, x) into (x^*, x^*); the homomorphism $\mathfrak{a}_\beta\rightarrow C/\mathfrak{c}$ therefore takes $x = xa_\beta+xa_\beta'$ into the class of elements congruent (mod $\mathfrak{c}$) to (x^*, x^*); and the isomorphism $C/\mathfrak{c}\rightarrow B$ takes this class, which consists of the two elements (e^*, x^*), $(0^*, x^*)$, into the element x^* in B. Hence the induced homomorphism $\mathfrak{a}_\beta\rightarrow B$ takes each element of $\mathfrak{b}_\beta \subset \mathfrak{a}_\beta$ into its correspondent under the homomorphism $\mathfrak{b}_\beta\rightarrow B$. A similar result obtains in case $\mathfrak{c}$ is the principal ideal generated by the element $(0^*, e^*)$. Finally, if $\mathfrak{c}$ is the principal ideal generated by the element $(0^*, 0^*)$, the quotient ring $C/\mathfrak{c}$ is isomorphic to C. In this case, we can define a homomorphism $C\rightarrow C/\mathfrak{d}\longleftrightarrow B$ by choosing the ideal $\mathfrak{d}$ as the principal ideal generated by $(e^*, 0^*)$ or by $(0^*, e^*)$, thus obtaining, as in the two preceding cases, a homomorphism $\mathfrak{a}_\beta\rightarrow B$ which takes each element of $\mathfrak{b}_\beta \subset \mathfrak{a}_\beta$ into its correspondent under the homomorphism $\mathfrak{b}_\beta\rightarrow B$. From the hypothesis that there exist homomorphisms $\mathfrak{a}_\alpha\rightarrow B$ for all $\alpha<\beta$ we therefore conclude that there exists a homomorphism $\mathfrak{a}_\beta\rightarrow B$, with the precise properties indicated above. The principle of transfinite induction thus establishes the existence of a homomorphism $\mathfrak{a}_\beta\rightarrow B$ for $\beta<\omega$, such that, when $\alpha<\beta$,

this homomorphism takes each element of $\mathfrak{a}_\alpha$ into its correspondent under the homomorphism $\mathfrak{a}_\alpha \to B$. Since we have $A = S_{\beta<\omega}\mathfrak{a}_\beta$, we conclude finally that there exists a homomorphism $A \to B$ which carries each element of $\mathfrak{a}_\beta$ into its correspondent under the homomorphism $\mathfrak{a}_\beta \to B$. By Theorem 49 the ideal $\mathfrak{p}$ determined in A by this homomorphism is a prime ideal. If A contains an element a not equal to 0 or to e, we may suppose that a does not belong to $\mathfrak{p}$. For, if we take $a_2 = a$, we see that the subring $\mathfrak{a}_2$ consists of the four elements $0, a, a', e$, and hence that we can define the homomorphism $\mathfrak{a}_2 \to B$ by the relations $0 \to 0^*$, $a' \to 0^*$, $a \to e^*$, $e \to e^*$. It is now possible to treat the case of a Boolean ring without unit: we have only to imbed the given ring A in a ring with unit in accordance with Theorem 1 and to construct a homomorphism from the imbedding ring to a two-element Boolean ring B in such a manner that an element $a \neq 0$ in A has as its correspondent in B the element e^*. This homomorphism evidently determines a homomorphism $A \to B$ and hence a prime ideal $\mathfrak{p}$ in A; and we may in particular require that $\mathfrak{p}$ shall not contain a specified element $a \neq 0$. It should be noted that the foregoing construction of a prime ideal consists essentially in the assignment, step-by-step, of the correspondents in B of the elements of A; and that at each step the construction reveals automatically whether the assignment is determined by the preceding steps or not. By particularizing the initial steps of the construction a little more, we can obtain a slightly sharper result. Let A be a Boolean ring with more than two elements and unit e; and let a and b be elements such that $a \neq e$, $b \neq e$, $ab \neq a$. We may then carry out the construction of the homomorphism $A \to B$ as before, setting $a_1 = e$, $a_2 = a$, $a_3 = b$ and determining the homomorphism $\mathfrak{a}_3 \to B$ as the correspondence which takes $0, b, a', a' \vee b, ab, a'b, a'b'$, and $ab \vee a'b'$ into 0^* and takes $a, e, b', a \vee b, a \vee b', a' \vee b', ab'$, and $a'b \vee ab'$ into e^*. We leave the details to the reader, noting that $\mathfrak{a}_3$ is in general a ring of sixteen elements which may reduce to one of eight or of four elements by virtue of equalities, such as $a' = b$, not incompatible with the relations $a \neq e$, $b \neq e$, $ab \neq a$. We then see that the homomorphism $A \to B$ takes a into e^* and b into 0^*, so that the resulting prime ideal $\mathfrak{p}$ contains b but not a. In the case of a Boolean ring A without unit, we obtain a similar result by imbedding A in a ring with unit, constructing a homomorphism from the imbedding ring to the two-element Boolean ring B so that the elements a and b in A with $ab \neq a$ are carried into e^* and 0^* respectively, and then determining $\mathfrak{p}$ from the restricted homomorphism $A \to B$.

In proving Theorem 63, we have obtained certain additional information concerning prime ideals. We now formulate this information as a separate theorem.

THEOREM 64. *If A is a Boolean ring containing elements a and b such that $ab \neq a$ or, equivalently, such that $a < b$ is false, then there exists a prime ideal $\mathfrak{p}$ in A which contains b and not a; and, if A is a Boolean ring containing ideals $\mathfrak{a}$ and $\mathfrak{b}$ such that $\mathfrak{b}$ is not a divisor of $\mathfrak{a}$, then there exists a prime ideal $\mathfrak{p}$ in A which is divisor of $\mathfrak{a}$ but not of $\mathfrak{b}$.*

We may restrict our reliance upon transfinite methods, however, by proving the following result:

THEOREM 65. *Theorem 64 follows from Theorem 63 without the intervention of transfinite methods or of the well-ordering hypothesis.*

We are now to assume that every Boolean ring containing two unequal elements contains a prime ideal. Let A be a Boolean ring containing elements a and b such that $ab \neq a$. Then the ideal $\mathfrak{c} = \mathfrak{a}(b) \vee \mathfrak{a}'(a)$ cannot contain $\mathfrak{a}(a)$, and, in particular, cannot coincide with the ideal $\mathfrak{e}$: for the relation $\mathfrak{a}(a) \subset \mathfrak{c}$ would imply $\mathfrak{a}(a) = \mathfrak{a}(a)\mathfrak{c} = \mathfrak{a}(a)\mathfrak{a}(b) \subset \mathfrak{a}(b)$ and hence $a = ab$, contrary to hypothesis. Since $\mathfrak{c} \neq \mathfrak{e}$, the quotient ring $A/\mathfrak{c}$ contains two or more elements and hence contains a prime ideal $\mathfrak{q}$. The elements of A which are carried by the homomorphism $A \rightarrow A/\mathfrak{c}$ into elements of $\mathfrak{q}$ constitute a prime ideal $\mathfrak{p}$ in accordance with Theorem 46. It is evident that $\mathfrak{p}$ is a divisor of $\mathfrak{c}$ and hence also of $\mathfrak{a}(b)$. On the other hand, $\mathfrak{p}$ is not a divisor of $\mathfrak{a}(a)$: for, if $\mathfrak{p} \supset \mathfrak{a}(a)$, we should have $\mathfrak{p} \supset \mathfrak{a}(a) \vee \mathfrak{c} = \mathfrak{a}(a) \vee \mathfrak{a}(b) \vee \mathfrak{a}'(a) = \mathfrak{e}$, $\mathfrak{p} = \mathfrak{e}$. Thus $\mathfrak{p}$ contains b but not a. Now let A be a Boolean ring containing ideals $\mathfrak{a}$ and $\mathfrak{b}$ where $\mathfrak{b}$ is not a divisor of $\mathfrak{a}$. The quotient ring $A/\mathfrak{b}$ contains an ideal $\mathfrak{c}$ consisting of all the images of elements of $\mathfrak{a}$ under the homomorphism $A \rightarrow A/\mathfrak{b}$, as we showed in Theorem 46. Since $\mathfrak{b}$ is not a divisor of $\mathfrak{a}$, the ideal $\mathfrak{c}$ contains an element other than the zero element. If we apply the first part of the theorem to the Boolean ring $A/\mathfrak{b}$, its zero element, and the indicated element of $\mathfrak{c}$, we see that $A/\mathfrak{b}$ contains a prime ideal $\mathfrak{q}$ which does not contain the specified element of $\mathfrak{c}$ and hence does not contain $\mathfrak{c}$. The elements of A which have images in $\mathfrak{q}$ under the homomorphism $A \rightarrow A/\mathfrak{b}$ thus constitute a prime ideal $\mathfrak{p}$ which contains $\mathfrak{b}$ but not $\mathfrak{a}$, in accordance with Theorem 46.

It is now easy to establish the Fundamental Proposition of Ideal Arithmetic for Boolean rings.

THEOREM 66. *In a Boolean ring A, every ideal other than $\mathfrak{e}$ is the product of all its prime ideal divisors. This result follows from Theorem 64 without the intervention of transfinite methods or of the well-ordering hypothesis.*

If $\mathfrak{a}$ is an ideal in a Boolean ring A such that $\mathfrak{a} \neq \mathfrak{e}$, then Theorem 64 shows that there exists a prime ideal which is a divisor of $\mathfrak{a}$ (but not of $\mathfrak{e}$), since $\mathfrak{e} \not\subset \mathfrak{a}$.

Hence the product of all the prime ideal divisors of $\mathfrak{a}$ exists and is an ideal divisor $\mathfrak{b}$ of $\mathfrak{a}$. If $\mathfrak{a}$ were not also a divisor of $\mathfrak{b}$, Theorem 64 would establish the existence of a prime ideal containing $\mathfrak{a}$ but not $\mathfrak{b}$, contrary to the definition of $\mathfrak{b}$. The relations $\mathfrak{a} \subset \mathfrak{b}$, $\mathfrak{b} \subset \mathfrak{a}$ together imply that $\mathfrak{a} = \mathfrak{b}$, as we wished to prove.

3. **The perfect representation.** We shall now turn to the study of the representation problem. We can give the fundamental theorem of the theory without any further preliminaries.

THEOREM 67. *Let A be a Boolean ring, $\mathfrak{a}$ an arbitrary ideal in A, $\mathfrak{E}$ the class of all prime ideals in A, $\mathfrak{I}$ the algebraic system of all ideals in A under the operations of unrestricted addition and finite multiplication, $\mathfrak{E}(\mathfrak{a})$ the class of all prime ideals which are not divisors of $\mathfrak{a}$, and $I(A)$ the algebraic system with the classes $\mathfrak{E}(\mathfrak{a})$ as elements and with the operations of forming unrestricted unions and finite intersections. Then the correspondence $\mathfrak{a} \rightarrow \mathfrak{E}(\mathfrak{a})$ determines an isomorphism $\mathfrak{I} \longleftrightarrow I(A)$ in accordance with the relations*

(1) $\mathfrak{E}(\mathfrak{a}) = \mathfrak{E}(\mathfrak{b})$ *if and only if* $\mathfrak{a} = \mathfrak{b}$;

(2) *if $\mathfrak{B}$ is any non-void class of ideals, then*

$$\mathfrak{E}\left(\sum_{\mathfrak{a} \in \mathfrak{B}} \mathfrak{a}\right) = \sum_{\mathfrak{a} \in \mathfrak{B}} \mathfrak{E}(\mathfrak{a});$$

(3) $\mathfrak{E}(\mathfrak{a}\mathfrak{b}) = \mathfrak{E}(\mathfrak{a})\mathfrak{E}(\mathfrak{b})$.

Let $\mathfrak{E}(a)$ denote the class $\mathfrak{E}(\mathfrak{a}(a))$ corresponding to the principal ideal $\mathfrak{a}(a)$; and let $B(A)$ be the algebraic system with the classes $\mathfrak{E}(a)$ as elements and with the operations of forming finite unions, symmetric differences (unions modulo 2), and finite intersections. Then $B(A)$ is a concrete Boolean ring or algebra of classes isomorphic to A by virtue of the correspondence $a \rightarrow \mathfrak{E}(a)$ in accordance with the relations

$$\mathfrak{E}(a) = \mathfrak{E}(b) \text{ if and only if } a = b; \tag{4}$$

$$\mathfrak{E}(a + b) = \mathfrak{E}(a) \,\Delta\, \mathfrak{E}(b); \tag{5}$$

$$\mathfrak{E}(a \vee b) = \mathfrak{E}(a) \subset \mathfrak{E}(b); \tag{6}$$

$$\mathfrak{E}(ab) = \mathfrak{E}(a)\mathfrak{E}(b). \tag{7}$$

The system $B(A)$ is a perfect reduced algebra of classes.

We first establish the numbered relations. Theorem 66 shows at once that (1) is valid. The relation (2) is equivalent to the assertion, obviously true, that a prime ideal fails to contain $S_{\mathfrak{a} \in \mathfrak{B}} \mathfrak{a}$ if and only if it fails to contain at least one of the ideals $\mathfrak{a}$ in $\mathfrak{B}$. Relation (3) is equivalent to the assertion that a prime ideal fails to contain $\mathfrak{a}\mathfrak{b}$ if and only if it contains neither $\mathfrak{a}$ nor $\mathfrak{b}$, and hence to the assertion, already proved in Theorem 40, that a prime ideal contains $\mathfrak{a}\mathfrak{b}$ if and only if it contains at least one of the ideals $\mathfrak{a}$ and $\mathfrak{b}$. The relations (4), (6), and (7) then follow immediately from (1), (2), and (3) respec-

tively by virtue of the respective relations (1), (3), and (4) of Theorem 31. The relation (5) is then an immediate consequence of (6) and (7) through the relations

$$\mathfrak{E}(a+b)\cup\mathfrak{E}(a)\mathfrak{E}(b) = \mathfrak{E}((a+b)\vee ab) = \mathfrak{E}(a\vee b) = \mathfrak{E}(a)\cup\mathfrak{E}(b),$$
$$\mathfrak{E}(a+b)\mathfrak{E}(a)\mathfrak{E}(b) = \mathfrak{E}((a+b)ab) = \mathfrak{E}(0) = \mathfrak{O},$$

where $\mathfrak{O}$ is the void subclass of $\mathfrak{E}$. The fact that the correspondences $\mathfrak{a}\rightarrow\mathfrak{E}(\mathfrak{a})$, $a\rightarrow\mathfrak{E}(a)$ set up isomorphisms $\mathfrak{J}\longleftrightarrow I(A)$, $A\longleftrightarrow B(A)$ respectively is now evident from the numbered relations. It remains to show that $B(A)$ is a perfect reduced algebra of classes. If $\mathfrak{p}$ is any prime ideal in A, then the classes $\mathfrak{E}(a)$ where a is in $\mathfrak{p}$ constitute a prime ideal in $B(A)$; and

$$\sum_{a\in\mathfrak{p}}\mathfrak{E}(a) = \mathfrak{E}(\mathrm{S}_{a\in\mathfrak{p}}\,\mathfrak{a}(a)) = \mathfrak{E}(\mathfrak{p})$$

in accordance with the relations (2); and, conversely, if the classes $\mathfrak{E}(a)$ constitute a prime ideal in $B(A)$, the corresponding elements a in A constitute a prime ideal $\mathfrak{p}$ in A. The class $\mathfrak{E}(\mathfrak{p})$, where $\mathfrak{p}$ is a prime ideal, clearly consists of all prime ideals other than $\mathfrak{p}$; and its complement $\mathfrak{E}'(\mathfrak{p})$ thus consists of $\mathfrak{p}$ alone. The classes $\mathfrak{E}(b)$ where b is not in $\mathfrak{p}$ obviously have the ideal $\mathfrak{p}$ in common. If b is any such element and a is an arbitrary element in $\mathfrak{p}$, then $b+ab$ cannot belong to $\mathfrak{p}$ since it is congruent (mod $\mathfrak{p}$) to b; hence $\mathfrak{E}(b+ab)$ contains $\mathfrak{p}$ and, by virtue of the relations $\mathfrak{E}(a)\mathfrak{E}(b+ab)=\mathfrak{E}(a(b+ab))=\mathfrak{E}(0)=\mathfrak{O}$, is contained in the complement of $\mathfrak{E}(a)$. Thus we see that the intersection of the classes $\mathfrak{E}(b)$ where b is not in $\mathfrak{p}$ is contained in the class

$$\prod_{a\in\mathfrak{p}}\mathfrak{E}'(a) = \Big(\sum_{a\in\mathfrak{p}}\mathfrak{E}(a)\Big)' = \mathfrak{E}'(\mathfrak{p}),$$

and thus consists of the ideal $\mathfrak{p}$ alone. These facts show that $B(A)$ is a reduced algebra of classes in accordance with Definition 10; and a perfect algebra of classes in accordance with Definition 12, Theorem 59 and Theorem 66.

We shall now introduce the following definition, justified by the result just established:

DEFINITION 13. *The algebra of classes $B(A)$ associated with a Boolean ring A by Theorem 67 is called the perfect representation of A.*

We call particular attention to the fact that, if A is a Boolean ring of just one element, the class $\mathfrak{E}$ is void and $B(A)$ is an algebra consisting of the void class alone.

The application of Theorem 56 to the perfect representation opens the way to the determination of all representations of a given Boolean ring. We obtain the following result:

THEOREM 68. *Let A, $\mathfrak{E}$, $\mathfrak{E}(\mathfrak{a})$, $I(A)$, $\mathfrak{E}(a)$, and $B(A)$ have the meanings specified in Theorem 67. Let $\mathfrak{H}$ be an arbitrary subclass of $\mathfrak{E}$; $\mathfrak{a}(\mathfrak{H}')$ the ideal consisting of all elements a such that $\mathfrak{E}(a) \subset \mathfrak{H}'$; $I(A, \mathfrak{H})$ the algebraic system of all classes $\mathfrak{H}\mathfrak{E}(\mathfrak{a})$ under the operations of forming unrestricted unions and finite intersections; and $B(A, \mathfrak{H})$ the algebraic system of all classes $\mathfrak{H}\mathfrak{E}(a)$ under the operations of forming finite unions, symmetric differences, and finite intersections. Then the correspondence $\mathfrak{E}(\mathfrak{a}) \to \mathfrak{H}\mathfrak{E}(\mathfrak{a})$ determines homomorphisms $I(A) \to I(A, \mathfrak{H})$, $I(A, \mathfrak{a}(\mathfrak{H}')) \to I(A, \mathfrak{H})$, the latter of which is an isomorphism if and only if $\mathfrak{E}(\mathfrak{a}(\mathfrak{H}')) = \mathfrak{H}'$ or, equivalently, $\mathfrak{H} = \mathfrak{E}'(\mathfrak{a}(\mathfrak{H}'))$. Similarly the correspondence $\mathfrak{E}(a) \to \mathfrak{H}\mathfrak{E}(a)$ determines a homomorphism $B(A) \to B(A, \mathfrak{H})$ and an isomorphism $B(A, \mathfrak{H}) \longleftrightarrow B(A/\mathfrak{a}(\mathfrak{H}'))$. The algebra of classes $B(A, \mathfrak{H})$ is perfect if and only if $\mathfrak{E}(\mathfrak{a}(\mathfrak{H}')) = \mathfrak{H}'$; and, when this condition is satisfied, $B(A, \mathfrak{H})$ is equivalent to $B(A/\mathfrak{a}(\mathfrak{H}'))$. If $\mathfrak{b}$ is an arbitrary ideal, then we have in particular the result that $B(A/\mathfrak{b})$ is equivalent to $B(A, \mathfrak{E}'(\mathfrak{b}))$. The ideal $\mathfrak{a}(\mathfrak{H}')$ is equal to $\mathfrak{e}$ when $\mathfrak{H}$ is void and to the product of the prime ideals in $\mathfrak{H}$ otherwise.*

Theorem 56 establishes the homomorphism $B(A) \to B(A, \mathfrak{H})$ at once; and an argument similar to that given in the proof of Theorem 56 shows that $I(A) \to I(A, \mathfrak{H})$. Since the correspondence $a \to \mathfrak{E}(a) \to \mathfrak{H}\mathfrak{E}(a)$ defines a homomorphism $A \to B(A, \mathfrak{H})$ by virtue of Theorem 67 and the result just proved, we see that there exists an ideal $\mathfrak{b}$ in A such that $A/\mathfrak{b} \longleftrightarrow B(A, \mathfrak{H})$ in accordance with Theorem 43. Since $a \to \mathfrak{E}(a) \to \mathfrak{O}$ if and only if $\mathfrak{H}\mathfrak{E}(a) = \mathfrak{O}$ or, equivalently, $\mathfrak{E}(a) \subset \mathfrak{H}'$, we see that this ideal coincides with the class $\mathfrak{a}(\mathfrak{H}')$ described above. Now Theorems 46 and 48 give detailed information concerning the prime ideals in the Boolean rings A, $A/\mathfrak{b}$, and $B(A, \mathfrak{H})$ under the indicated correspondence. If $\mathfrak{p}$ is a prime ideal divisor of $\mathfrak{b}$, then its image $\mathfrak{p}^*$ in $A/\mathfrak{b}$ is a prime ideal, and the image of $\mathfrak{p}^*$ in $B(A, \mathfrak{H})$ is also a prime ideal; conversely, a prime ideal in $B(A, \mathfrak{H})$ is the image of a prime ideal $\mathfrak{p}^*$ in $A/\mathfrak{b}$, and $\mathfrak{p}^*$ is in turn the image of a unique prime ideal divisor of $\mathfrak{b}$. Thus we see that the class of prime ideals in $B(A, \mathfrak{H})$ is in biunivocal correspondence with the class $\mathfrak{E}^*$ of prime ideals in $A/\mathfrak{b}$ and also with the class $\mathfrak{E}'(\mathfrak{b})$ of those prime ideals in A which divide $\mathfrak{b}$. If an ideal $\mathfrak{a}$ in A has the image $\mathfrak{a}^*$ in $A/\mathfrak{b}$, then the prime ideal divisors of $\mathfrak{a}^*$ are in biunivocal correspondence with those prime ideal divisors of $\mathfrak{a}$ which contain $\mathfrak{b}$. Hence the biunivocal correspondence between $\mathfrak{E}'(\mathfrak{b})$ and $\mathfrak{E}^*$ takes $\mathfrak{E}'(\mathfrak{b})\mathfrak{E}(\mathfrak{a})$ into $\mathfrak{E}^*(\mathfrak{a}^*)$; and $\mathfrak{E}'(\mathfrak{b})\mathfrak{E}(\mathfrak{a}_1) = \mathfrak{E}'(\mathfrak{b})\mathfrak{E}(\mathfrak{a}_2)$ whenever $\mathfrak{a}_1$ and $\mathfrak{a}_2$ have the same image in $A/\mathfrak{b}$. It follows that the correspondence $a \to \mathfrak{E}(a) \to \mathfrak{H}\mathfrak{E}(a)$ determines a correspondence $\mathfrak{E}^*(\mathfrak{a}^*) \to \mathfrak{E}'(\mathfrak{b})\mathfrak{E}(\mathfrak{a}) \to \mathfrak{H}\mathfrak{E}'(\mathfrak{b})\mathfrak{E}(\mathfrak{a}) = \mathfrak{H}\mathfrak{E}(\mathfrak{a})$ which is a homomorphism $I(A/\mathfrak{b}) \to I(A, \mathfrak{H})$. We have here used the obvious fact that $\mathfrak{E}(\mathfrak{b}) \subset \mathfrak{H}'$ to write $\mathfrak{H}\mathfrak{E}'(\mathfrak{b}) = \mathfrak{H}$. Since the principal ideal $\mathfrak{a}^*$ in $A/\mathfrak{b}$ is the image of at least one principal ideal $\mathfrak{a}(a)$ in A, we may specialize

this homomorphism to the systems $B(A/\mathfrak{b})$ and $B(A, \mathfrak{H})$. We shall now show that it becomes an isomorphism $B(A/\mathfrak{b}) \longleftrightarrow B(A, \mathfrak{H})$ when so specialized. In order that the principal ideals $\mathfrak{a}(a)$ and $\mathfrak{a}(b)$ in A have the same image in $A/\mathfrak{b}$ we must have $\mathfrak{a}(a) \vee \mathfrak{b} = \mathfrak{a}(b) \vee \mathfrak{b}$, hence $a \vee c_1 = b \vee c_2$ where c_1 and c_2 are in $\mathfrak{b}$, and hence $a \equiv b \pmod{\mathfrak{b}}$, by Theorems 46, 16, and 44 respectively; and the converse is also true. On the other hand, $\mathfrak{H}\mathfrak{E}(a) = \mathfrak{H}\mathfrak{E}(b)$ is equivalent to $\mathfrak{H}\mathfrak{E}(a+b) = \mathfrak{H}\mathfrak{E}(a)\Delta\mathfrak{H}\mathfrak{E}(b) = \mathfrak{O}$ and hence to $\mathfrak{E}(a+b) \subset \mathfrak{H}'$, $a+b \in \mathfrak{b}$, and $a \equiv b \pmod{\mathfrak{b}}$. It follows immediately that the asserted relation of isomorphism holds. In the case where $\mathfrak{H} = \mathfrak{E}'(\mathfrak{b})$ or, equivalently, $\mathfrak{E}(\mathfrak{b}) = \mathfrak{E}(\mathfrak{a}(\mathfrak{H}')) = \mathfrak{H}'$, it is evident that the correspondence $\mathfrak{E}^*(\mathfrak{a}^*) \rightarrow \mathfrak{H}\mathfrak{E}(\mathfrak{a}) = \mathfrak{E}'(\mathfrak{b})\mathfrak{E}(\mathfrak{a})$ yields an isomorphism $I(A/\mathfrak{b}) \longleftrightarrow I(A, \mathfrak{H})$. Furthermore, it is evident that the correspondence between $\mathfrak{E}^*$ and $\mathfrak{H} = \mathfrak{E}'(\mathfrak{b})$ renders the two algebras of classes $B(A/\mathfrak{b})$ and $B(A, \mathfrak{H})$ equivalent. Thus $B(A, \mathfrak{H})$ is a perfect reduced algebra of classes in this case, in accordance with Theorems 67 and 55. On the other hand, if $\mathfrak{E}(\mathfrak{b}) \neq \mathfrak{H}'$, there exists a prime ideal $\mathfrak{p}$ in A which belongs to $\mathfrak{H}'$ but not to $\mathfrak{E}(\mathfrak{b})$. It follows that $\mathfrak{p} \supset \mathfrak{b}$ and $\mathfrak{E}(\mathfrak{p}) \supset \mathfrak{H}$. Since $\mathfrak{p}$ is a divisor of $\mathfrak{b}$, its correspondent $\mathfrak{p}^*$ in $A/\mathfrak{b}$ is a prime ideal; and the classes $\mathfrak{H}\mathfrak{E}(a)$ where a is in $\mathfrak{p}$ constitute a prime ideal in $B(A, \mathfrak{H})$. Now the homomorphism $I(A/\mathfrak{B}) \rightarrow I(A, \mathfrak{H})$ takes $\mathfrak{E}^*(\mathfrak{p}^*)$ into $\mathfrak{H}\mathfrak{E}(\mathfrak{p}) = \mathfrak{H}$ and therefore cannot be an isomorphism since it also takes $\mathfrak{E}^* = \mathfrak{E}^*(\mathfrak{e}^*)$ into $\mathfrak{H}$. Furthermore the union of the classes $\mathfrak{H}\mathfrak{E}(a)$, where a is in $\mathfrak{p}$, is given by

$$\sum_{a \in \mathfrak{p}} \mathfrak{H}\mathfrak{E}(a) = \mathfrak{H} \sum_{a \in \mathfrak{p}} \mathfrak{E}(\mathfrak{a}(a)) = \mathfrak{H}\mathfrak{E}(\mathrm{S}\, \mathfrak{a}(a)) = \mathfrak{H}\mathfrak{E}(\mathfrak{p}) = \mathfrak{H}.$$

The algebra of classes $B(A, \mathfrak{H})$ therefore fails to have property (2) of Theorem 59, and cannot be perfect. Finally we prove that $\mathfrak{b} = \mathfrak{a}(\mathfrak{H}')$ is characterized in the manner indicated in the statement of the theorem. If $\mathfrak{H}$ is void, then $\mathfrak{E}(a) \subset \mathfrak{H}'$ for every a and $\mathfrak{a}(\mathfrak{H}')$ is obviously equal to $\mathfrak{e}$. If $\mathfrak{H}$ is not void, then the product of all the prime ideals in $\mathfrak{H}$ is an ideal $\mathfrak{c}$. Since $\mathfrak{E}(\mathfrak{p}) \supset \mathfrak{a}(\mathfrak{H}')$ when $\mathfrak{p}$ is in $\mathfrak{H}$, it is evident that $\mathfrak{p} \supset \mathfrak{a}(\mathfrak{H}')$ when $\mathfrak{p}$ is in $\mathfrak{H}$, and hence that $\mathfrak{c} \supset \mathfrak{a}(\mathfrak{H}')$. On the other hand, if a is an element of $\mathfrak{c}$, we have $\mathfrak{a}(a) \subset \mathfrak{c} \subset \mathfrak{p}$ for every $\mathfrak{p}$ in $\mathfrak{H}$ or, equivalently, $\mathfrak{E}(a) \subset \mathfrak{E}(\mathfrak{p})$ for every $\mathfrak{p}$ in $\mathfrak{H}$. Since $\mathfrak{E}(\mathfrak{p})$ does not contain $\mathfrak{p}$, it follows that $\mathfrak{E}(a) \subset \mathfrak{H}'$ and that a is in $\mathfrak{a}(\mathfrak{H}')$. Thus we have $\mathfrak{c} \subset \mathfrak{a}(\mathfrak{H}')$ as well as $\mathfrak{a}(\mathfrak{H}') \subset \mathfrak{c}$, and conclude that $\mathfrak{c} = \mathfrak{a}(\mathfrak{H}')$ as desired. This result shows in particular that, if $\mathfrak{H} = \mathfrak{E}'(\mathfrak{b})$ where $\mathfrak{b}$ is an arbitrary ideal in A, then $\mathfrak{a}(\mathfrak{H}') = \mathfrak{b}$.

We now complete the theory of representations by means of the following result:

THEOREM 69. *If B is an algebra of classes homomorphic to a Boolean ring A and if $\mathfrak{b}$ is the ideal in A determined by the homomorphism $A \rightarrow B$, then there*

exists a class $\mathfrak{H}$ of prime ideals in A related to $\mathfrak{b}$ through the equation $\mathfrak{a}(\mathfrak{H}') = \mathfrak{b}$ such that B is equivalent to $B(A, \mathfrak{H})$. In order that B be perfect it is necessary and sufficient that $\mathfrak{H} = \mathfrak{E}'(\mathfrak{b})$. The only perfect algebras of classes isomorphic to A are those equivalent to $B(A)$.

The prime ideals in B are in biunivocal correspondence with the prime ideal divisors of $\mathfrak{b}$ in A, by Theorem 48. Those prime ideals in B which determine one-element subclasses of the basic class E of B in accordance with Theorem 57 (9) thus define a subclass $\mathfrak{H}$ of $\mathfrak{E}'(\mathfrak{b})$. This correspondence between E and $\mathfrak{H}$ is biunivocal and defines the equivalence between B and $B(A, \mathfrak{H})$: for if a is any element of A, the prime ideals which contain it and also contain $\mathfrak{b}$ constitute the class $\mathfrak{E}'(\mathfrak{b})\mathfrak{E}'(a)$; these prime ideals are carried by the homomorphism $A \rightarrow B$ into the prime ideals in B which contain the image of a; of the latter ideals those and only those which determine one-element classes disjoint from the image of a can correspond to ideals in $\mathfrak{H}$; and the indicated correspondence between $\mathfrak{H}$ and E thus takes $\mathfrak{H}\mathfrak{E}'(a)$ into the complement of the image of a, $\mathfrak{H}\mathfrak{E}(a)$ into the image of a. Now those elements a which are taken by the homomorphism $A \rightarrow B$ into the void class in B are precisely those for which $\mathfrak{H}\mathfrak{E}(a)$ is void, as we see from the foregoing remarks. Thus it follows that the ideal $\mathfrak{b}$ determined by the homomorphism in question coincides with the ideal $\mathfrak{a}(\mathfrak{H}')$. If B is perfect, then every prime ideal in B determines a one-element class by Theorem 59; and it follows that $\mathfrak{H} = \mathfrak{E}'(\mathfrak{b})$. On the other hand, if $\mathfrak{H} = \mathfrak{E}'(\mathfrak{b})$, then $B(A, \mathfrak{H})$ is perfect by Theorem 68; and B, being equivalent to $B(A, \mathfrak{H})$, is obviously perfect also. If B is isomorphic to A, then $\mathfrak{b} = \mathfrak{o}$. Since $\mathfrak{H} = \mathfrak{E}'(\mathfrak{b}) = \mathfrak{E}'(\mathfrak{o}) = \mathfrak{E}$ in this case, B is thus equivalent to $B(A, \mathfrak{E})$, and hence to $B(A)$, if it is perfect. Since $B(A)$ is isomorphic to A and is perfect, any algebra of classes equivalent to $B(A)$ also has these properties.

We may bring the discussion to a close by formulating in precise terms our remarks in §1 concerning the relation between the representation theory and the Fundamental Proposition of Ideal Arithmetic. We have

THEOREM 70. *The following propositions are equivalent without the use of transfinite methods or the well-ordering hypothesis*:

(1) *every Boolean ring possesses an isomorphic algebra of classes*;

(2) *the Fundamental Proposition of Ideal Arithmetic is valid in every Boolean ring.*

In Theorem 67, we have shown that (2) implies (1), without the use of transfinite arguments. In Theorem 58 (4) we showed without the use of such arguments that (1) implies the existence of a prime ideal in any Boolean ring

which is a reduced algebra of classes with a basic class E which is non-void; and hence that (1) implies the Fundamental Existence Proposition, proved as Theorem 63 by other methods. The work of §2 shows that Theorem 63 implies (2), without the use of transfinite arguments, as indicated in Theorem 66.

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

SIMPLE ALGEBRAS OF DEGREE p^e OVER A CENTRUM OF CHARACTERISTIC p*

BY

A. ADRIAN ALBERT

1. **Introduction.** The study of normal simple algebras A over a field F has been reduced to the case where the degree of A is a power of a prime p. The theory then splits sharply into two cases distinguished by the hypothesis that the characteristic of F is or is not p. We shall restrict our attention to the former case.

It is well known that every normal simple algebra A over F is similar to a crossed product B. But this result is of little aid in a study of A, since in fact the degree of B is in general not a power of p. We shall prove here, however, that in our case every A is *similar to a cyclic algebra whose degree is a power of* p.

If K is obtained from F by adjoining the p^eth roots of quantities of F to F, for fixed e, then K is said to have exponent e over F. We shall show that the *exponent* of an *algebra* A is p^e where e is the exponent of the above K of least exponent, which splits A. Moreover A has exponent p^e only if A is similar to a direct product of *cyclic division algebras* D_i whose *exponents and degrees are equal to* $p^{e_i} \leqq p^e$, D_1 of degree p^e.

2. **Commutative division algebras over** F. Let F be an infinite field of characteristic $p \neq 0$ and K be a commutative division algebra of order n over F. A quantity k of K is called separable† or inseparable according as its minimum equation has not or has multiple roots. We shall also say that K is separable or inseparable according as K does not or does contain inseparable quantities.

Every separable K is a simple algebraic extension $Z=F(\alpha)$ of F, where α is separable of degree n over F. Conversely if α is separable so is any polynomial in α, so that $Z=F(\alpha)$ is separable. We notice that then every subfield of Z is separable. Moreover if $F(\alpha^p) \neq F(\alpha)=Z$, then $\alpha^p=v$ has the property that $Z=Z_1(\alpha)$, $\alpha^p=v$ in Z_2. But the equation $\alpha^p=v$ is inseparable. Hence we have

---

* Presented to the Society, September 13, 1935; received by the editors July 22, 1935.

† For these definitions and elementary properties see B. L. van der Waerden, *Moderne Algebra*, vol. I. I believe they are due to E. Steinitz. See also Deuring's *Algebren*, Springer, 1935, for references to the numerous concepts used here. These are in the algebraic part of Deuring's book and the references are principally to work of Dickson, Wedderburn, Hasse, Brauer, Noether, and myself.

THEOREM 1. *Let $Z=F(x)$ be separable of degree n over F. Then $Z=F(x^p)$.*

The largest separable sub-field K_0 of an inseparable field K is a field whose degree n_0 is called* the reduced degree of K. Every quantity α of K has the property

$$\alpha^{p^e} = \alpha_0 \text{ in } K_0. \tag{1}$$

If moreover $\alpha^{p^{e-1}}$ is not in K_0 for some α of K we shall call e the *exponent* of K. The field K is obviously obtained by adjoining certain p^eth roots of quantities of K_0 to K_0. If $K_0=F$ *we shall call K a Kummer field over F.* We now prove

THEOREM 2. *Let Z be a separable field of degree m over a Kummer field K of exponent e over F, and Z_0 be the largest sub-field over F of Z, Z contain K. Then Z is the direct product*

$$Z = Z_0 \times K, \tag{2}$$

Z_0 has degree m over F, Z_0 is the field of all quantities α^{p^e} of Z, α in Z.

For $Z=K(\xi)$ where ξ is a root of a separable equation

$$f(x) = x^m + a_1x^{m-1} + \cdots + a_m = 0 \qquad (a_i \text{ in } K), \tag{3}$$

irreducible in K. Thus $\eta=\xi^{p^e}$ is a root of

$$g(y) = y^m + b_1y^{m-1} + \cdots + b_m = 0 \qquad (b_i \text{ in } F), \tag{4}$$

where $b_i=a_i^{p^e}$. By Theorem 1 we have $Z=K(\eta)$, so that $g(y)$ is irreducible in K. But then $Z_0=F(\eta)$ has degree n over F and (2) holds. The largest separable sub-field of Z must contain Z_0 and is a field $F(\zeta)=F(\zeta^{p^e})$ where ζ^{p^e} is obviously in Z_0. Hence Z_0 is the largest separable sub-field of Z.

The fields Z and Z_0 of Theorem 2 are equivalent under the correspondence (*not an isomorphism over F*)

$$\alpha \longleftrightarrow \alpha^{p^e} = \alpha_0 \qquad (\alpha \text{ in } Z). \tag{5}$$

If $\alpha\neq\beta$ then $\alpha_0\neq\beta_0$ since $\alpha^{p^e}-\beta^{p^e}=(\alpha-\beta)^{p^e}=0$ only when $\alpha=\beta$. Hence we have immediately

THEOREM 3. *Let Z and Z_0 be defined as in Theorem 2. Then Z is normal over K with automorphisms*

$$S: \qquad \alpha \longleftrightarrow \alpha^S$$

if and only if Z_0 is normal over F with automorphisms

$$S: \qquad \alpha_0 \longleftrightarrow \alpha_0^S, \qquad \alpha_0 = \alpha^{p^e}, \qquad \alpha_0 = (\alpha^S)^{p^e}.$$

* See van der Waerden, loc. cit.

COROLLARY I. *The field Z is cyclic over K if and only if Z_0 is cyclic over F.*

Let K be a simple Kummer extension of exponent e over F, so that $K=K_e=F(y)$, $y^{p^e}=\gamma$ in F. Then

$$K_e > K_{e-1} > \cdots > K_1 > K_0 = F, \tag{6}$$

where

$$y_i = y^{p^{e-i}}, \qquad K_i = F(y_i). \tag{7}$$

The field K_i has degree p^i over F, degree p over K_{i-1}, and $1, y_i, \cdots, y_i^{p-1}$ are linearly independent in K_{i-1}. In particular every quantity k of K has the form

$$k = k_0 + k_1 y + \cdots + k_{p-1}y^{p-1} \qquad (k_i \text{ in } K_{e-1}), \tag{8}$$

and, if $q=p^{e-1}$, then

$$k^q = k_0 + k_1^q y_1 + \cdots + k_{p-1}^q y_1^{p-1} \qquad (k_i^q \text{ in } F) \tag{9}$$

is in F if and only if $k_1=k_2=\cdots=k_{p-1}=0$. Thus we have proved

THEOREM 4. *A quantity k of K_e generates K_e if and only if k is not in K_{e-1}.*

We shall require

THEOREM 5. *Let Z be separable of degree m over $K=K_e=F(y)$, $y^{p^e}=\gamma$ in F. Then there exists a quantity c in Z such that*

$$w = N_{Z/K}(c) \tag{10}$$

has the property $K=F(w)$, that is, w generates K.

For $Z=Z_0\times K$, $Z_0=F(x)$, x is a root of

$$f(x) = \sum_{i=0}^{p-1} a_i(x^p)x^i = 0. \tag{11}$$

The polynomials $a_i(x^p)$ have coefficients in F and at least one $a_i(x^p)\neq 0$ for $i=r\neq 0$, since x is separable. Hence $a_r(\lambda^p y^p)$ is a polynomial in y^p with coefficients which are polynomials in λ with coefficients in F. These latter polynomials are not all identically zero, and thus there exists a λ_0 in F such that $a_r(\lambda_0^p y^p)\neq 0$,

$$f(\lambda_0 y) = \sum_{i=0}^{p-1} b_i y^i, \qquad b_i = a_i(\lambda_0^p y^p)\lambda^i \text{ in } K_{e-1}. \tag{12}$$

The quantity $f(\lambda_0 y)$ is not in K_{e-1} and, if $c=x-\lambda_0 y$, then

$$g(c) = f(c+\lambda_0 y) = c^m + B_1 c^{m-1} + \cdots + B_{m-1}c + f(\lambda_0 v) = 0. \tag{13}$$

But

$$N_{Z/K}(c) = (-1)^m f(\lambda_0 y) = w$$

is in K_e and not in K_{e-1}. By Theorem 4 we have $K = F(c)$.

The field K_e has degree p over $K_{e-1} = F(y^p)$ and, if δ is in K_{e-1}, and w is as above, then δw is not in K_{e-1}. This gives

THEOREM 6. *Let Z be as in Theorem 5 and δ be in K. Then there exists a quantity c of Z such that*

$$\delta_0 = N_{Z/K}(c)\delta \tag{14}$$

generates K, $K = F(\delta_0)$.

3. **Matrices with elements in F.** The theorems of the author's paper* *On normal simple algebras* hold for an arbitrary infinite field F although they were derived for the case where F is non-modular. The extension to the more general case may be made by the elementary considerations of the next three sections.

The matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & 1 \\ \alpha_n & \alpha_{n-1} & \alpha_{n-2} & \cdots & \alpha_1 \end{pmatrix}, \qquad \alpha_i \text{ in } F, \tag{15}$$

has

$$|\lambda I - A| = \phi(\lambda) = \lambda^n - \alpha_1\lambda^{n-1} - \cdots - \alpha_n = 0 \tag{16}$$

as its characteristic equation. It may be easily verified that the identity matrix $I, A, A^2, \cdots, A^{n-1}$ are linearly independent in any scalar field containing F and that $\phi(\lambda) = 0$ is actually the minimum equation of A. Moreover every n-rowed square matrix with $\phi(\lambda) = 0$ as *minimum* equation is similar in F to A.

Every matrix B with elements in F is similar to a matrix†

$$\begin{pmatrix} A_1 & & \\ & \cdot & \\ & & A_r \end{pmatrix} \tag{17}$$

where A_i is a matrix whose characteristic equation is its minimum equation. The characteristic determinants $D_i(\lambda)$ of the A_i are called the invariant fac-

* These Transactions, vol. 34 (1932), pp. 620–625.

† These are the classical results on matrices with elements in F.

tors of B and $D_{i-1}(\lambda)$ is divisible by $D_i(\lambda)$, $i=2, \cdots, \nu$. Thus $D_1(A)=0$, and $D_1(\lambda)=0$ is in fact the minimum equation of the matrix B. It is evident that a matrix $\overline{B}$ is similar to B if and only if $\overline{B}$ has the same invariant factors and hence the same canonical form given by (15), (17) as B.

When $D_1(\lambda)$ is irreducible we must have every $D_i(\lambda) \equiv D_1(\lambda)$, $n=m\nu$ where A_1 has m rows. Thus

$$B = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_1 \end{pmatrix}, \tag{18}$$

and $\overline{B}$ is similar to B if and only if $\overline{B}$ has the same *irreducible* minimum equation as B.

4. **Sub-fields of M_n.** Let M_n be the algebra of all n-rowed square matrices with elements in F. If Z is a *separable* sub-field of M_n and $\overline{Z}$ is equivalent to Z, then $Z=F(B)$, $\overline{Z}=F(\overline{B})$, where B and $\overline{B}$ have the same irreducible minimum equation. Hence $B=P\overline{B}P^{-1}$ where P is a non-singular quantity of M_n, and the inner automorphism of M_n, defined by the transformation of its quantities by P, carries B into $\overline{B}$. If $CB=BC$ then

$$C = (C_{ij}) \qquad (i, j = 1, \cdots, \nu) \tag{19}$$

and $C_{ij}A_1=A_1C_{ij}$. But it is known* that then C_{ij} is a polynomial in A_1 with coefficients in F. Hence the algebra of all quantities of M_n commutative with every quantity of Z is equivalent to a total matric algebra of degree ν over Z, $n=m\nu$, Z of degree m over F.

The above result cannot be extended to inseparable sub-fields K of M_n without further argument since the known proofs that $C_{ij}A_1=A_1C_{ij}$ if and only if C_{ij} is in $F(A_1)$ depend upon the hypothesis that A_1 has its characteristic roots all distinct. Hence we must treat this case here. We first prove

LEMMA 1. *Let $F(y)$ be a sub-field of M_p, $y^p=\gamma$ in F, F an infinite field of characteristic p. Then the only quantities of M_p commutative with y are quantities of $F(y)$.*

For we have seen that we may take

$$y = \begin{pmatrix} 0 & 1 & 0 \cdots 0 \\ 0 & 0 & 1 \cdots 0 \\ \cdot & \cdot & \cdots\cdots \\ 0 & 0 & 0 \cdots 1 \\ \gamma & 0 & 0 \cdots 0 \end{pmatrix}. \tag{20}$$

* Cf. the proof in Wedderburn's *Lectures on Matrices*, pp. 26–27.

Since F is infinite there exists a ξ in F such that $\xi \neq 0, 1, \cdots, p-1$. The matrix

$$x = \begin{pmatrix} \xi & & & \\ & \xi+1 & & \\ & & \ddots & \\ & & & \xi+p-1 \end{pmatrix} \tag{21}$$

is non-singular and has the property $yx=(x+1)y$. An elementary computation then gives

$$e_{11} = \frac{[x-(\xi+1)][x-(\xi+2)]\cdots[x-(\xi+p-1)]}{1\cdot 2\cdots(p-1)} \tag{22}$$

where e_{ij} is the matrix with unity in the ith row and jth column and zeros elsewhere. Moreover it is well known* that

$$e_{ij} = y^{1-i}e_{11}y^{j-1}. \tag{23}$$

But the relation $yx=(x+1)y$ then implies that every quantity of M_p is a linear combination of the p^2 quantities

$$x^i y^j \qquad (i, j = 0, 1, \cdots, p-1). \tag{24}$$

The algebra M_p has order p^2 over F and hence (24) are linearly independent in F. Thus $1, y, \cdots, y^{p-1}$ are *left linearly independent* in the commutative algebra $F(x)$. This latter algebra is not a field, however.

We now write $z=\sum_{i=0}^{p-1} a_i(x)y^i$ and have

$$zy - yz = \sum_{i=0}^{p-1} [a_i(x+1) - a_i(x)]y^{i+1} = 0, \tag{25}$$

so that $a_i(x)=a_i(x+1)$, $i=0, \cdots, p-1$. But then the element of the first row of $a_i(x)$ is $a_i(\xi)=a_i(\xi+1)$ so that $a_i(\xi)=a_i(\xi+1)=a_i(\xi+2)=\cdots=a_i(\xi+p-1)$ and $a_i(x)$ is a scalar matrix. Hence z is in $F(y)$.

We now prove

THEOREM 7. *Let Z be any sub-field of degree m of M_n. Then the sub-algebra of M_n commutative with Z is a total matric algebra M_ν of degree $\nu=nm^{-1}$ over Z, and any sub-field $\overline{Z}$ of M_n equivalent to Z may be carried into Z by an inner automorphism of M_n.*

The theorem is trivial when $n=1$. Hence assume it true for total algebras M_ν of degree $\nu<n$.

Let Z_0 be the largest separable sub-field of Z. Then we have proved the

* Cf. these Transactions, vol. 33 (1931), pp. 690–711, formula (20) on p. 702. Formula (18) was incorrectly printed there and should read $e_{11}y^{k-1}=e_{1k}$.

above theorem for Z_0 and hence may carry the sub-field $\overline{Z}_0$ of $\overline{Z}$ into Z_0. Hence we take $Z_0=\overline{Z}_0$ and have proved that the algebra of all quantities of M_n commutative with Z_0 of degree m_0 over F is a total matric algebra of degree $n_0=nm_0^{-1}$ over Z_0. Every inner automorphism of M_{n_0} over Z_0 is an inner automorphism of M_n leaving Z_0 invariant and M_{n_0} contains both Z and $\overline{Z}$. If $Z_0\neq F$ then $n_0<n$ and we may carry Z into $\overline{Z}$ by an inner automorphism of M_{n_0}; the algebra of all quantities of M_{n_0} commutative with Z is a total matric algebra of degree $n_0m_1^{-1}$ over Z. But Z has degree m_1 over Z_0, $m=m_1m_0$, $n_0m_1^{-1}=nm^{-1}=\nu$ and the algebra of all quantities of M_n commutative with Z is M_ν over Z. Hence let $Z_0=F$, Z be a Kummer field of degree $m=p\mu$ over F. Then Z has degree μ over $K=F(y)$, $y^p=\gamma$ in F.

The minimum equation of y is irreducible in F and $\overline{Z}>\overline{K}=F(\bar{y})$, $\bar{y}^p=\gamma$, $\bar{y}$ may be carried into y by an inner automorphism of M_n. Thus we take $y=\bar{y}$, $K=\overline{K}$, and y in the form (18) with A_1 given by (20). By Lemma 1 the sub-algebra of M_n commutative with y is M_{ν_0} of degree $\nu_0=np^{-1}$ over K. Moreover M_{ν_0} contains Z, $\overline{Z}$, and our induction states that there exists an inner automorphism of M_{ν_0} carrying $\overline{Z}$ into Z, the sub-algebra of M_{ν_0} commutative with Z is M_ν over Z, $\nu=\nu_0\mu^{-1}=nm^{-1}$. This inner automorphism of M_{ν_0} is an inner automorphism of M_n, and M_ν is obviously the sub-algebra of M_n commutative with Z. This proves Theorem 7.

5. **Simple algebras over F.** Let A be a simple algebra of degree n over its centrum F and let Z be a sub-field of degree m of A. If Z_0 is the maximum separable sub-field of Z then the sub-algebra B_0 over Z_0 of all quantities of A commutative with every quantity of Z_0 is well known to be a simple algebra of degree $\nu_0=nm_0^{-1}$ over its centrum Z_0 of degree m_0 over F. But we may in fact prove

THEOREM 8. *The sub-algebra of A commutative with Z is a normal simple algebra B of degree nm^{-1} over its centrum Z. Moreover if $\overline{Z}$ is a scalar field equivalent to Z then*

$$A_{\overline{Z}} = M_m \times \overline{B}, \tag{26}$$

where $\overline{B}$ over $\overline{Z}$ is equivalent to B over Z.

For it is obviously sufficient to prove the above theorem when $Z_0=F$. There exists a separable splitting field X of A and $A_X=M_n$ over X. But the composite $Z_X=(Z, X)$ is evidently an inseparable field of degree n over X and Theorem 6 states that the sub-algebra of A_X commutative with Z_X is M_ν over Z_X. Evidently $M_\nu=B_X$ so that B is normal simple of degree ν over Z. The remainder of the proof is as in the non-modular case quoted* in §3.

* These Transactions, vol. 34 (1932), pp. 620–625.

6. **Kummer splitting fields and cyclic algebras.** Let F have characteristic p and Ω be a perfect extension of F. Then if A is any normal simple algebra of degree p^e over F the algebra A_Ω is a total matric algebra.* The field Ω may be taken to consist of quantities δ such that δ^{p^s} is in F for some s depending on δ, and thus every sub-field of Ω of finite degree over F is a Kummer field over F. In particular the sub-field K of Ω, which contains all of the coefficients in the expression of an ordinary matric basis of A_Ω in terms of the basal units of A, has finite degree over F and splits A. We have proved

THEOREM 9. *Every normal simple algebra A of p^e over F of characteristic p has a Kummer splitting field.*

We next prove

THEOREM 10. *If there exist no cyclic fields of degree p over F then there exist no normal division algebras of degree p^e over F.*

For let D be a normal division algebra of degree p^e over F. Theorem 9 states that there exists a Kummer field K such that K does not split D but $K(y)$, $y^p=\gamma$ in K, does split D. Thus D_K is similar to a normal division algebra E of degree p over K which is split by $K(y)$. The author has then proved† that $E=(\gamma, Z, S)$ where Z is cyclic of degree p over K. By Theorem 2, $Z=Z_0\times K$ where Z_0 is cyclic of degree p over F.

If F possesses no cyclic extensions of degree p over F, then Theorem 10 implies that every normal simple algebra of degree p^e over F is a total matric algebra. This case is then complete, so we shall henceforth assume that there exist cyclic fields of degree p over F.

The author has shown‡ that our above assumption then implies that there exist cyclic fields Z_f of degree p^f over F for every f. Moreover if Z_f is given then there exists a cyclic Z_e, $e>f$, such that Z_e contains Z_f.

In particular let $f=e-1$. Then the author has proved† that

$$Z_e = F(x), \qquad x^p = x + a + \lambda \qquad (a \text{ in } Z_{e-1}, \lambda \text{ in } F), \tag{27}$$

with generating automorphism S given by that of Z_e and

$$x^S = x + \beta. \tag{28}$$

The quantities a, β are uniquely determined quantities of Z_{e-1}, and the author

* These Transactions, vol. 36 (1934), pp. 388–394.

† See the paper *On normal division algebras of degree p^e over F of characteristic p*, these Transactions, vol. 39 (1936), pp. 183–188.

‡ Bulletin of the American Mathematical Society, vol. 40 (1934), pp. 625–631.

has moreover proved that if a, β, λ are given, $x^p = x + a + \lambda$ is irreducible in Z_{e-1}, $F(x) = Z_e$ is cyclic of degree p^e over F.

Our hypothesis that there exist cyclic fields of degree p over F then implies that there exist cyclic algebras of arbitrary degree p^e over F. The author has proved*

LEMMA 2. *Let (γ, Z, S) be cyclic of degree $n = p^e$ over F and Y be the cyclic sub-field of Z of degree p^{e-1}. Then*

$$(\gamma, Z, S)^p = M_n^{p-1} \times M_p \times (\gamma, Y, S). \tag{29}$$

As an immediate application of Lemma 2 we obtain†

THEOREM 11. *Let $A = (\gamma, Z, S)$ be cyclic of degree p^e over F of characteristic p. Then Z is contained in a cyclic field $\overline{Z}$ of degree p^{e+1} over F, and the algebra*

$$B = (\gamma, \overline{Z}, S) \tag{30}$$

has the property

$$A \sim B^p. \tag{31}$$

7. **Theory of cyclic representations.** A field F of characteristic p is said to be perfect if there are no Kummer fields of degree p over F. Then there are no normal division algebras of degree p^e over F. This, combined with our previous discussion, shows that the theory of normal simple algebras of degree p^e over F becomes trivial except when there exist both Kummer and cyclic fields of degree p over F. We may therefore restrict our attention to this remaining‡ non-trivial case, and have seen that there now exist cyclic and Kummer fields of arbitrary degree p^e over F.

A total matric algebra of degree p over F contains fields isomorphic to any fields of degree p over F. Every normal simple algebra A of degree p over F is either a total matric algebra or a normal division algebra, and the author has shown§ that in the latter case A is a cyclic algebra if and only if it has a simple Kummer sub-field of degree p over F. Thus we have the case $e = 1$ of

* See the American Journal of Mathematics, vol. 54 (1932), pp. 1–13, for my proof holding for any field.

† Theorem 11 is false when F does not have characteristic p, since in fact the algebra $A = (1, i, j, ij)$, $i^2 = j^2 = -1$, $ji = -ij$, over the field R of rational numbers, is not similar to B^2 for any B of degree 4 over its centrum R. Note also that $A \sim B$, for normal simple algebras A and B, means $A = M \times D$, $B = M_0 \times D$ where D is a division algebra, M and M_0 are total matric algebras.

‡ We have actually assumed that the field F does not possess a certain type of algebraic closure. Such fields F of course exist and there do exist normal division algebras of degree p^e over some fields F.

§ These Transactions, vol. 39, loc. cit.

THEOREM 12. *A normal simple algebra A of degree p^e over F is cyclic if and only if A has a simple Kummer sub-field $K=F(y)$ of degree p^e over F.*

For we now assume the above theorem true for algebras of degree $p, p^2, \cdots, p^{e-1}$ and write $y_0=y^{p^{e-1}}$, $y_0{}^p=\gamma$ in F, $K=F(y)$ of degree p^e over F. The sub-algebra of A commutative with y_0 is a normal simple algebra C_{y_0} of degree p^{e-1} over $K_0=F(y_0)$ by Theorem 8. The algebra C_{y_0} contains the simple Kummer field $K_0(y)$ of degree p^{e-1} over K_0 and the hypothesis of our induction states that C_{y_0} is a cyclic algebra (g, X, S) where X is cyclic of degree p^{e-1} over K_0 and y is in K_0. If d is any quantity of X and

$$g' = N_{X/K_0}(d)g, \tag{32}$$

then $C_{y_0}=(g', X, S)$. Hence Theorem 6 states that we may choose g' to generate K_0. Obviously $C_{y_0}=C_{y'}$ so that we may write $C_{y_0}=(y_0, X, S)$,

$$yx = x^S y \qquad (x \text{ in } X), \tag{33}$$

without loss of generality.

Theorem 3 states that $X=X_0\times K_0$ where X_0 is cyclic of degree p^{e-1} over F and

$$yx_0 = x_0{}^S y \qquad (x_0 \text{ in } X_0). \tag{34}$$

The algebra C_{X_0} of all quantities of A commutative with every x_0 of X_0 is a normal simple algebra of degree p over F by Theorem 8. If C_{X_0} is a total matric algebra over X_0, then C_{X_0} contains fields equivalent to any field of degree p over X_0. But then A contains cyclic fields of degree p^e over F with X_0 as sub-field and is a cyclic algebra. Hence let C_{X_0} be a division algebra.

The algebra C_{X_0} contains the sub-field $K_0=F(y_0)$, $y_0{}^p=\gamma$, and the author has then proved* that $C_{X_0}=(\gamma, Z, S)$ where $Z=X_0(z)$ is cyclic of degree p over X_0,

$$z^p = z + a, \qquad y_0 z = (x_0 + \epsilon)y_0 \qquad (a \text{ in } X_0). \tag{35}$$

Also we may take

$$\epsilon = (-1)^{e-1} = T_{X_0/F}(\beta), \tag{36}$$

where β is given as in (28) and $T_{X_0/F}$ is the trace function. The quantity y transforms every quantity of X_0 into a quantity of X_0 and hence every quantity of C_{X_0} into a quantity of C_{X_0}. We thus write

$$yy_0 = y_0 y, \qquad yzy^{-1} = z_y = \sum_{i=0}^{p-1} b_i y_0{}^i \qquad (b_i \text{ in } Z). \tag{37}$$

* These Transactions, vol. 39, loc. cit.

But obviously $y_0 z_y = (z_y + \epsilon) y_0$ so that we have

$$\sum_{i=0}^{p-1} b_i(z+\epsilon) y_0^{i+1} = \epsilon y_0 + \sum_{i=0}^{p-1} b_i(z) y_0^{i+1}. \tag{38}$$

Thus $b_i(z+\epsilon) = b_i$ is in X_0 for $i = 1, \cdots, m$, while $b_0(x_0+\epsilon) = b_0 + \epsilon$. Then $b_0 = kz + \delta$ with k an integer and δ in X_0, $k(z+\epsilon)+\delta = kz + k\epsilon + \delta = kz + \delta + \epsilon$, $k = 1$,

$$z_y = z + P, \qquad P \text{ in } X = X_0(y_0). \tag{39}$$

The transformation of z by powers of y then gives $q = p^{e-1}$, $y^q = y_0$,

$$y^q z y^{-q} = z + \epsilon = z + T_{X/K}(P), \tag{40}$$

so that $T_{X/K}(P - \beta) = 0$. The author has then demonstrated* the existence of quantity Q in X such that

$$\beta = P + Q^S - Q. \tag{41}$$

We write

$$x_e = z + Q, \tag{42}$$

and have $y x_e y^{-1} = z + P + Q^S = x_e + \beta$. Thus $y_0 x_e y_0^{-1} = x_e + \epsilon$ satisfies a normed equation $x_e{}^p = x_e + a_e$ where $(x_e+\beta)^p = x_e + \beta + a_e{}^S$,

$$a_e{}^S - a_e = \beta^p - \beta. \tag{43}$$

But then $F(x_e)$ is cyclic of degree p^e over F and A is a cyclic algebra.

Conversely let $A = (\gamma, Z, S)$ be cyclic of degree p^e over F. From Lemma 2, $A = M \times (\delta, Y, S)$ where M is a total matric algebra and Y is the sub-field of Z of degree $q = p^f$ over F, $\delta \neq \lambda^p$ for any λ of F. The algebra (δ, Y, S) thus has a simple Kummer sub-field $F(y_0)$, $y_0{}^q = \delta$, and $M \times F(y_0)$ contains a simple Kummer field $F(y)$, $y^{p^e} = \gamma$, of degree p^e over F.

As a consequence of Theorem 12 we have

THEOREM 13. *Let A and B be cyclic algebras whose degrees are powers of the characteristic p of F. Then $A \times B$ is a cyclic algebra.*

For Theorem 12 states that B has a maximal simple Kummer sub-field $K_e = F(y)$. We write $A_K = (y_1, Z, S)$ where Z is cyclic of degree p over K. By Theorem 6 we may take y_1 to generate K and thus A_K has a simple Kummer sub-field $K_0 = K(y_0)$, $y_0{}^{p^e} = y_1$. Obviously $K_0 = F(y_0)$, $y_0{}^q = \delta$ in F, K_0 has degree $q = p^{e+f}$ over F. But the degree of $A \times B$ is q and $A \times B$ contains K_0. Hence Theorem 12 states that $A \times B$ is cyclic.

* These Transactions, vol. 39, loc. cit.

A further consequence of Theorem 12 is given by

THEOREM 14. *A normal division algebra D of degree p^f over F is similar to a cyclic algebra if and only if D has a Kummer splitting field of degree p^e over F.*

For by Theorem 12 algebra D has a Kummer splitting field of degree p^e over F when D is similar to a cyclic algebra.

Conversely let K split D. If $K=F(y)$ is simple of degree p^e over F, then $e>f$ and Theorem 12 states that D is similar to a cyclic algebra of degree p^e over F. Let then $K=F(y_1, \cdots, y_r)$ of degree p^e be a splitting field of D and assume as the basis of an induction on r that our theorem is true for algebras with Kummer splitting fields with at most $r-1$ generators. We may also assume that no sub-field of K splits D. Then $K_0=F(y_1)$ does not split D and has degree p^k over F, $K=K_0(y_2, \cdots, y_r)$ has degree p^{e-k} over K_0. Thus D is similar to a normal simple algebra A of degree p^e over F with K as maximal sub-field. The algebra B of all quantities of A commutative with y is a normal simple algebra of degree p^{e-k} over K_0 and is split by $K=K_0(y_2, \cdots, y_r)$. By the hypothesis of our induction B is a cyclic algebra (γ, Z, S) over K_0 with γ in K_0. Theorem 6 implies that we may take γ to generate K_0, $\gamma=y_1$, and thus B contains a sub-field $K(y)$, $y^{p^{e-1}}=y_1$, $y^{p^e}=\delta$ in F. By Theorem 12 algebra A is a cyclic algebra.

As an immediate corollary of Theorems 13 and 9 we have our principal result:

THEOREM 15. *Every normal division algebra of degree p^f over F of characteristic p is similar to a cyclic algebra of degree $p^e \geqq p^f$.*

8. **Theory of exponents.** A cyclic field Z of degree $\nu=p^n$ over F has the generation $Z=F(x)$ where x is a root of a (separable) cyclic equation of degree ν over F. Theorem 1 states that any quantity z of Z has the form

$$z = \sum_{i=1}^{p} \alpha_i x^{p(i-1)}, \qquad \alpha_i \text{ in } F. \tag{44}$$

As a consequence of (44) we may prove

THEOREM 16. *The exponent $\rho=p^\phi$ of a normal division algebra D of degree p^e over F is the least integer p^f such that f is the exponent of a Kummer splitting field of D.*

For D is similar to a cyclic algebra (γ, Z, S) of degree $\nu=p^n$ over F. Let K be a Kummer splitting field of D of least exponent f over F so that D_K is a total matric algebra. Then $\gamma=N_{Z_0/K}(z_0)$, where z_0 is in $Z_0=Z_K$. The

quantity $z=z_0^q$ is in Z if $q=p^f$ and hence $\gamma^q=N_{Z/F}(z)$. It is well known* that then the exponent $\rho=p^\phi$ is at most p^f, $\phi\leqq f$.

Conversely let $\rho=p^\phi$. If $\phi=1$ then $\gamma^p=N_{Z/F}(z)$ where z in Z has the form (44). Write

$$\beta_i = \alpha_i^{1/p}. \tag{45}$$

Not every β_i is in F, for otherwise $z_1=\sum_{i=1}^{p}\beta_i x^{i-1}$ is in Z, $z_0=z_1$, $\gamma^p=[N_{Z/F}(z_1)]^p$, γ must be the norm of z_1 when F has characteristic p. Then D is a total matric algebra contrary to hypothesis. But $K=F(\beta_1, \cdots, \beta_r)$ has exponent unity over F, $z=z_1^p$ where z_1 is in Z_K, $\gamma=N_{Z_1/K}(z_1)$, K splits D. Thus $f=\phi=1$.

We have proved Theorem 16 true for $\phi=f=1$. Assume it true for algebras of exponents p, p^2, $\cdots$, $p^{\phi-1}$, and let D have exponent p^ϕ. The algebra D^p has exponent $p^{\phi-1}$ and the hypothesis of our induction implies that D^p has a Kummer splitting field H of exponent $\phi-1$. The algebra $(D_H)^p$ is a total matric algebra so that D_H has exponent 1 or p. In either case D_H has a Kummer splitting field $K\geqq H$ of exponent at most unity over H, and K has exponent $\psi\leqq\phi$ over F. Obviously K splits D so that $\phi\geqq\psi\geqq f$. But we have proved that $\phi\leqq f$ so that $\phi=f$ and our theorem is proved.

Let K be a Kummer field of degree p^r and exponent unity over F. Then

$$K=K_r > K_{r-1} > \cdots > K_1 > K_0 = F, \tag{46}$$

where

$$K_i = F(\alpha_1^{1/p}, \cdots, \alpha_i^{1/p}) = K_{i-1}(\alpha_i^{1/p}), \quad \alpha_i \text{ in } F, \qquad (i=1, \cdots, r). \tag{47}$$

Obviously $K_r\neq F(\beta_1^{1/p}, \cdots, \beta_s^{1/p})$ for any β_j in F and $s<r$. We now prove

THEOREM 17. *Let A be a normal simple algebra of exponent p over F and p^r be the minimum degree of all Kummer splitting fields of A of exponent unity. Then A is similar to an algebra*

$$D_1 \times D_2 \times \cdots \times D_r, \tag{48}$$

where D_i is a cyclic division algebra of degree p over F.

For we may assume that A is a normal simple algebra of degree p^r over F with K as a maximal sub-field. If $r=1$ then $K=F(\alpha^{1/p})$ splits A and the author has shown† that then A is a cyclic algebra (γ, Z, S) of degree p over F.

We make an induction on r and let $H=K_{r-1}$. By hypothesis H does not split A, so that, by Theorem 8, the sub-algebra B of A of all quantities com-

* American Journal of Mathematics, loc. cit.

† These Transactions, vol. 39, loc. cit.

mutative with every quantity of H is a normal division algebra of degree p over H. By our above case $r=1$, $B=(\alpha_1, Z, S)$ over H and, since $Z=Z_0\times H$, algebra $B=(\alpha_1, Z_0, S)\times H=D_1\times H$ where D_1 is cyclic of degree p over F. Thus $A=D_1\times A_1$ where A_1 is a normal simple algebra of degree p^{r-1} with H as a maximal sub-field. If r_0 is the least integer for which A_1 has a Kummer splitting field of degree p^{r_0} over F and exponent unity, then the corresponding integer for A is at most r_0+1. Thus $r\leqq r_0+1$. But evidently $r_0\leqq r-1$, $r\geqq r_0+1$, $r_0=r-1$. By our induction $A_1=D_2\times\cdots\times D_r$, A has the form (48) as desired.

We shall finally prove

THEOREM 18. *Every normal simple algebra of exponent p^e is similar to a direct product of cyclic normal division algebras*

$$D_1\times\cdots\times D_t, \tag{49}$$

where the exponent and degree of D_i are equal and at most p^e, and where D_1 has exponent p^e.

For we have the above result when $e=1$. Let us then make an induction on e. The algebra A of exponent p^e is such that A^p has exponent p^{e-1}. By the assumption of our induction $A^p\sim B_1\times\cdots\times B_r$, where B_i is a cyclic normal division algebra of degree and exponent $p^{f_i}\leqq p^{e-1}$. If $p^{f_i}<p^{e-1}$ for all values of i, then the exponent of $B_1\times\cdots\times B_r$ is at most p^{e-2}, a contradiction. Hence we may assume that the exponent of B_1 is p^{e-1}. By Theorem 11 there exist cyclic normal division algebras D_i of degree p^{f_i+1} such that $D_i^p\sim B_i$. Evidently D_i has exponent p^{f_i+1}. Moreover

$$A^p\sim A_0^p,$$

where

$$A_0=D_1\times\cdots\times D_r.$$

But if A_0^{-1} is reciprocal to A_0, the algebra $(AA_0^{-1})^p$ is a total matric algebra. Either AA_0^{-1} is a total matric algebra and A is similar to A_0, or AA_0^{-1} has exponent p. In the latter case $AA_0^{-1}\sim D_{r+1}\times\cdots\times D_t$ where D_{r+j} has exponent and degree p. Then

$$A\sim D_1\times\cdots\times D_t,$$

as desired.

The above canonical form for A evidently yields an algebra of exponent p^e if and only if algebra B of exponent p given by $q=p^{e-1}$, $B=D_1^q\times\cdots\times D_t^q$, is not a total matric algebra. Hence in (49) we assume that $D_1, \cdots, D_s$ have

degree and exponent p^e while $D_{s+1}, \cdots, D_t$ have degree at most p^{e-1}. We write

$$D_i = (\gamma_i, Z_i, S_i), \qquad \gamma_i \text{ in } F,$$

where Z_i is cyclic of degree p^e over F. Then it is known* that the degree of D_i is its exponent if and only if

$$B_i = (\gamma_i, Y_i, S)$$

is a division algebra, where Y_i is the unique sub-field of degree p of Z_i. Thus A has exponent p^e if and only if $B_1 \times \cdots \times B_s$ is not a total matric algebra.

9. **A conjecture.** The structure of a normal division algebra relative to exponent has been in doubt since the author proved† the existence of *primary* normal division algebras of degree eight and exponent four. The results of §8 completely solve the problem for normal division algebras of degree p^e over F of characteristic p. The only remaining case is that of division algebras of degree p^e over F of characteristic not p, since the study of normal division algebras has been reduced to the case where the degree and exponent are a power of a prime. For this case we may hope to prove the

CONJECTURAL THEOREM. *A normal division algebra D has exponent p^e only if D is similar to a direct product of normal division algebras*

$$D_1 \times \cdots \times D_t,$$

where the exponent of D_i is its degree $p^{e_i} \leqq p^e$, and D_1 has degree p^e.

Thus the author's example of an algebra D of degree eight and exponent four *may possibly be* that of a normal division algebra D obtained by

$$M_2 \times D = D_1 \times D_2,$$

where D_1 and D_2 are normal division algebras of degree and exponent four, M_2 is a total matric algebra of degree two.

The author has not attempted to prove the above conjecture, but its proof is probably very difficult. It would be an important advance in the theory, however.

---

* American Journal of Mathematics, loc. cit.

† Bulletin of the American Mathematical Society, vol. 39 (1933), pp. 265–277.

THE UNIVERSITY OF CHICAGO,
CHICAGO, ILL.

ASYMPTOTIC SOLUTIONS OF CERTAIN LINEAR DIFFERENTIAL EQUATIONS IN WHICH THE COEFFICIENT OF THE PARAMETER MAY HAVE A ZERO*

BY

HENRY SCHEFFÉ

CHAPTER I. THE GIVEN EQUATION

1. Introduction. Equations of the type

$$\frac{d^n u}{dz^n} + p_1(z)\frac{d^{n-1}u}{dz^{n-1}} + p_2(z)\frac{d^{n-2}u}{dz^{n-2}} + \cdots + \{p_n(z) - \rho^n\phi^n(z)\}u = 0$$

have been studied by various mathematicians, including Birkhoff,† Schlesinger, and Horn, with regard to the asymptotic forms of the solutions for large complex values of the parameter ρ. The second-order case goes back to Liouville. The essential differences between the present treatment and previous treatments of the nth-order case are the following:

(1) Before this it has been necessary to assume the coefficient $\phi^n(z)$ to be bounded from zero. We permit it to have a zero of any integral order ν.

(2) The variable z has been confined to a finite interval of the real axis. We allow it to vary over complex values in a finite region.

(3) In order that the forms be valid for all ρ in the neighborhood of $\rho = \infty$, it has been necessary to require that the coefficient $\phi^n(z)$ be real, except possibly for a constant imaginary factor. We dispense with this restriction.

Besides proving the existence of solutions asymptotic to n simple independent expressions, we consider the following problem arising from the first of the above generalizations: What is the form of a solution determined by boundary values at the point where the coefficient $\phi^n(z)$ has its zero? At a point where $\phi^n(z)$ is different from zero the problem becomes elementary, since it is always possible to find a non-singular matrix of simple forms for n solutions and their first $n-1$ derivatives which is valid in some neighborhood including the point. But in our case the simple forms are *not* valid at the zero. The problem is complicated by the *Stokes phenomenon*, that is, the discontinuity, as a solution is taken across a "Stokes line," of the coefficients

* Presented to the Society, April 20, 1935; received by the editors July 3, 1935.

† Birkhoff, *On the asymptotic character*, etc., these Transactions, vol. 9 (1908), p. 219, which contains references to the work of the other mathematicians mentioned above.

(hereafter called "Stokes multipliers") that multiply the simple forms in the asymptotic expression of the solution.

The case for $n=2$ has been studied by Langer.* Applications of his work not only have led to new results for some of the classical equations of physics (Bessel,* Weber,† Hermite,† Mathieu‡) but are of importance in quantum physics.§ Mathematically, perhaps the most alluring aspect of the problem is the Stokes phenomenon. In going to the nth-order case it is then of interest to discover how the quantitative dependence of the Stokes multipliers on ν and how the location of the Stokes lines generalize.

Langer was able to obtain a "related equation" solvable by Bessel functions and thus utilize the classical knowledge of their asymptotic properties. Because of the apparent lack of a suitable standard equation of the nth order for which the Stokes multipliers for the principal solutions are known, a considerable part of the work lies in developing this necessary adjunct. A feature of interest is the way in which the multipliers are found, by use of the Stokes formula. In his letter‖ in the Abel memorial edition of the Acta Mathematica, Stokes forecasts the general applicability of this method of finding the multipliers for solutions appropriately expressible as power series with gamma function coefficients. The labor of calculating formal solutions at ∞ would be lessened in such cases, since the formula yields the leading terms, and the existence of true solutions asymptotic to the formal solutions could then be predicated upon one of the fundamental papers in that field.¶

2. **Canonical form.** We assume the equation to be treated has undergone preliminary transformations leaving it in normal form and transferring the zero of $\phi^n(z)$ to the origin. The *given equation*

---

* Langer, *On the asymptotic solutions*, etc., these Transactions, vol. 33 (1931), p. 23, and vol. 34 (1932), p. 447. In the former article, but not the latter, z is restricted to real values. The region of validity R_ν need not be finite and ν need only be positive, where $\phi^n(z)\equiv(z-z_0)^\nu\phi_1{}^n(z)$ and $\phi_1{}^n(z)$ is analytic and bounded from zero in R_ν. See also his symposium article, *The asymptotic solutions*, etc., Bulletin of the American Mathematical Society, vol. 40 (1934), p. 545, which contains further references to his work.

† N. Schwid, *The asymptotic forms of the Hermite and Weber functions*, these Transactions, vol. 37 (1935), p. 339.

‡ Langer, *The solutions of the Mathieu equation*, etc., these Transactions, vol. 36 (1934), p. 637.

§ Birkhoff, *Quantum mechanics and asymptotic series*, Bulletin of the American Mathematical Society, vol. 39 (1933), p. 681.

‖ Stokes, *On the discontinuity of arbitrary constants*, etc., Acta Mathematica, vol. 26 (1902), p. 393.

¶ Trjitzinsky, *Analytic theory of linear differential equations*, Acta Mathematica, vol. 62 (1933), p. 167, treats the subject with the greatest generality. The pioneer article is by Poincaré, *Sur les intégrales irrégulières des équations linéaires*, Acta Mathematica, vol. 8 (1886), p. 295. Important advances were made by Horn, Birkhoff.

$$(1) \qquad u^{(n)}(z, \rho) + 0 + p_2(z)u^{(n-2)}(z, \rho) + \cdots + \{p_n(z) - \rho^n\phi^n(z)\}u(z, \rho) = 0$$

is then subject to five hypotheses:

(i) *R_z is a finite, closed, simply connected region, containing $z=0$ as interior point.*

(ii) *The coefficients $\phi^n(z)$ and $p_j(z)$, $j=2, 3, \cdots, n$, are analytic in R_z.*

(iii) *The coefficient $\phi^n(z)$ is of the form $z^\nu\phi_1^n(z)$, where ν is a positive integer or zero, and $\phi_1^n(z)\neq 0$ in R_z.*

Hypothesis (iv), given in §5, implies a further restriction on the coefficient $\phi^n(z)$, while (v), which is most conveniently stated in §8, further qualifies the region R_z.

CHAPTER II. AN IMPORTANT SPECIAL CASE

3. **The t-equation: two sets of solutions.** The simplest possible equation of the type (1) is

$$\frac{d^n}{dx^n}y(x, \rho) - \rho^n x^\nu y(x, \rho) = 0,$$

and we shall consider this first. However, we shall find that the treatment of the general case can be based directly on the results for this special case.

It is preferable to study the simpler transform

$$(2) \qquad \frac{d^n}{dt^n}y(t) - t^\nu y(t) = 0,$$

which we shall designate the *t-equation*, and which we shall later force to serve as our standard equation of the nth order. The transformation to the t-equation is achieved by the substitution

$$x = \rho^{-n/p}t,$$

where

$$p \equiv n + \nu.$$

Power series for that set of solutions of (2) which is principal at the origin were obtained by an elegant method by M. H. Molins* in 1876; they are also easily found by the usual method. The series†

---

* Mémoires de l'Académie des Sciences, Inscriptions et Belles-Lettres de Toulouse, (7), vol. 8 (1876), p. 167.

† If this formula becomes indeterminate at $t=0$ we define $y_j^{(m)}(0) \equiv \lim_{t\to 0} y_j^{(m)}(t)$. We use the notation

$$(a)_m \equiv a(a-1)(a-2)\cdots(a-m+1) \qquad (m = 1, 2, 3, \cdots); \ (a)_0 \equiv 1.$$

Note $(j)_m=0$ if j is an integer $<m$.

$$y_j^{(m)}(t) \equiv \frac{t^{j-m}}{j!}\left[(j)_m + \sum_{l=1}^{\infty} \frac{(lp+j)_m}{\prod_{k=1}^{l}(kp+j)_n}\, t^{lp}\right] \qquad (j, m = 0, 1, \cdots, n-1)$$

for the principal solutions and their derivatives converge for all t.

For later use it is desirable to express the coefficients of these power series as quotients of gamma functions. It is also of subsequent advantage to study the solutions of the t-equation on a ξ-plane by means of the mapping

$$t \equiv \left(\frac{p}{n}\xi\right)^{n/p}. \tag{3}$$

In the ξ-plane it is necessary to construct a Riemann surface with branch point at the origin, appropriate to the single-valued representation of $\arg \xi$. It will suffice to employ the sector

$$-Q\frac{\pi}{n} \leqq \arg \xi \leqq Q\frac{\pi}{n}. \tag{4}$$

Q is an even integer whose definition is conveniently postponed to §5; we may remark, however, that there is always more than one sheet.

The functions*

$$y_j^{(m)}(t) = \alpha_{jm}\sigma_{jm}(\xi), \tag{5}$$

where

$$\sigma_{jm}(\xi) \equiv \xi^{n(j-m)/p}\sum_{l=0}^{\infty} \frac{\left(\frac{\xi}{n}\right)^{ln}}{\prod_{i=0}^{n-1}\Gamma(l+\lambda_{ijm}+1)},$$

$$\lambda_{ijm} \equiv \frac{j-i}{p} - \begin{cases} 1, & i < m, \\ 0, & i \geqq m, \end{cases}$$

$$\alpha_{jm} \equiv p^{(nj+\nu m)/p}n^{n(m-j)/p}\kappa_j,$$

* The following remarks apply to formulas (5) et seq.

(1) If $m>j$ one of the gamma functions in the denominator of the term for $l=0$ is infinite, and the term is then defined to be zero, thus removing the apparent singularity.

(2) In this and all subsequent formulas involving a real positive constant to a real power, the real positive value is to be used.

(3) For purposes of calculation it might be desirable when $\nu<n-1$ to replace the n gamma functions in the denominator by ν gamma functions of the same type in the numerator and a single gamma function in the denominator. This is possible by applying the multiplication theorem of Gauss. See Whittaker and Watson's *Modern Analysis*, §12·15, or Nielsen's *Handbuch der Theorie der Gammafunktion*, §6.

and

$$\kappa_j \equiv \frac{1}{j!}\prod_{i=0}^{n-1}\Gamma(\lambda_{ij0}+1),$$

are then single-valued* at every point of the ξ-surface (4). Furthermore, although under the transformation (3), corresponding to some point t there may be several points on the ξ-surface, all with the same absolute value but with arguments differing by integral multiples of $2\pi p/n$, the value of $y_j^{(m)}(t)$ calculated from the formula (5) is the same for all these points ξ.

The series (5) in ascending powers of ξ converge very slowly for large ξ. It is therefore natural to seek series in descending powers to satisfy the t-equation. Let $\omega_0, \omega_1, \cdots, \omega_{n-1}$ be the nth roots of unity, and define

$$\beta \equiv \frac{n-1}{2}.$$

Then by formally differentiating the series

$$s_k(t) \equiv \frac{e^{\omega_k \xi}}{\xi^{\nu\beta/p}}\left\{1+\frac{b_{k01}}{\xi}+\frac{b_{k02}}{\xi^2}+\cdots\right\} \qquad (k=0,1,\cdots,n-1)$$

with respect to t,

$$(6)\qquad s_k^{(m)}(t) = \omega_k^m\left(\frac{p}{n}\right)^{m\nu/p}\frac{e^{\omega_k\xi}}{\xi^{\nu(\beta-m)/p}}\left\{1+\frac{b_{km1}}{\xi}+\frac{b_{km2}}{\xi^2}+\cdots\right\}$$
$$(m=0,1,\cdots,n),$$

and substituting in the t-equation, one finds that the constants b_{k0l} can be successively determined and are unique. However, the series (6) are in general divergent and are hence called *formal solutions*.

Beginning with Poincaré's famous paper† in the Acta Mathematica in 1886, and culminating in a recent paper by Trjitzinsky‡ in the same journal, much progress has been made on the problem of showing that true solutions of a linear differential equation exist which under certain conditions are asymptotic to the formal solutions. The latter writer shows that for a very inclusive class of equations the neighborhood of ∞ can be completely covered by a finite number of regions, associated with each of which there exists a full set of true solutions uniformly asymptotic in the region to the formal solutions.

* It is to be understood here and hereafter that if arg Z is specified, the value to be chosen for arg Z^μ, μ real, is μ arg Z.

† Loc. cit.

‡ Loc. cit.

It now becomes necessary to digress for a moment in order to subdivide the ξ-surface. The rays L_τ from the origin,

$$L_\tau: \qquad \arg \xi = \theta_\tau \equiv \tau \frac{\pi}{n} \qquad (\tau = 0, \pm 1, \pm 2, \cdots, \pm Q),$$

divide it into $2Q$ sectors S_τ,

$$S_\tau: \qquad \theta_{\tau-1} \leqq \arg \xi \leqq \theta_\tau, \qquad -Q+1 \leqq \tau \leqq Q.$$

Although the system of S-sectors suffices for most of the discussion, a different system of sectors covering the ξ-surface is required to simplify the statement of our final results: Each of the rays L_{2q+1}, q integral, $-Q<2q+1<Q$, is inclosed in a sector $D_{q,q+1}$ whose central angle 2δ is small compared with π/n,

$$D_{q,q+1}: \qquad \theta_{2q+1} - \delta \leqq \arg \xi \leqq \theta_{2q+1} + \delta.$$

The region between two successive D-sectors, and having as its bisector the line L_{2q}, $q=0, \pm 1, \pm 2, \cdots, \pm(Q/2-1)$, is denoted by T_q,

$$T_q: \qquad \theta_{2q-1} + \delta \leqq \arg \xi \leqq \theta_{2q+1} - \delta.$$

The "left-over" regions are then called $T_{Q/2}$ and $T_{-Q/2}$,

$$T_{Q/2}: \qquad \theta_{Q-1} + \delta \leqq \arg \xi \leqq \theta_Q,$$

$$T_{-Q/2}: \qquad \theta_{-Q} \leqq \arg \xi \leqq \theta_{-Q+1} - \delta.$$

All the S, D, and T regions are closed except at $\xi = \infty$.

The fundamental theorem* of Trjitzinsky's paper applied to the t-equation insures the existence of true solutions and their t-derivatives, $\{{}_\tau\bar{y}_k^{(m)}(t)\}$, $k, m=0, 1, \cdots, n-1$, analytic in t, such that

$$(7) \qquad {}_\tau\bar{y}_k^{(m)}(t) \sim s_k^{(m)}(t)$$

uniformly with respect to ξ for ξ in S_τ, that is,

$${}_\tau\bar{y}_k^{(m)}(t) = \omega_k^m \left(\frac{p}{n}\right)^{m\nu/p} \frac{e^{\omega_k \xi}}{\xi^{\nu(\beta-m)/p}} \left\{1 + \frac{b_{km1}}{\xi} + \cdots + \frac{b_{k,m,\mu-1}}{\xi^{\mu-1}} + \frac{{}_\tau b_{km\mu}(\xi)}{\xi^\mu}\right\},$$

where $|{}_\tau b_{km\mu}(\xi)| < M_\mu$ for $|\xi| \geqq N$, ξ in S_τ, $\mu = 1, 2, 3, \cdots$.

4. **The Stokes multipliers.** Corresponding to a region S_τ, the subscripts k of the nth roots of unity $\{\omega_k\}$ can be arranged so that the following ordering of the real parts of $\{\omega_k \xi\}$ is maintained for ξ in S_τ:

* Loc. cit., p. 208. For $n>2$ our regions S_i are identical with Trjitzinsky's regions R_i and require no further subdivision into regions R_i' and R_i''. For $n=2$ the Trjitzinsky regions R_i consist each of two of our quadrants S_τ, but in this case subdivision is necessary into R_i' and R_i'', each of which may be taken as a quadrant S_τ.

$$R(\omega_0\xi) \geqq R(\omega_1\,\xi) \geqq \cdots \geqq R(\omega_{n-1}\xi).$$

We shall call ${}_\tau\bar{y}_0$ the *dominant* solution and ${}_\tau\bar{y}_1$ the *first sub-dominant* solution in S_τ. The actual values of ω_0 and ω_1 for each of the regions S_τ are given in the table at the end of this section.

Between the solutions $\{{}_\tau\bar{y}_k\}$ and the principal solutions $\{y_j\}$ a non-singular linear relationship must exist,

$$y_j = \sum_{k=0}^{n-1} {}_\tau c_{jk}\, {}_\tau\bar{y}_k. \tag{8}$$

It is of the greatest importance for us to evaluate the Stokes multipliers $\{{}_\tau c_{jk}\}$, or at least the multipliers ${}_\tau c_{j0}$ and ${}_\tau c_{j1}$ of the dominant and first sub-dominant solutions. To accomplish this we utilize* a formula of Stokes.

By an ingenious method Stokes showed† that a power series whose terms are of the type

$$X_l = \frac{\prod_{i=1}^{r} \Gamma(l + a_i + 1)}{\prod_{i=1}^{s} \Gamma(l + b_i + 1)}\, x^l, \quad \text{where} \quad \begin{cases} \mu \equiv s - r > 0, \\ a_i,\ b_i, \text{ real}, \\ a_i \neq -1, -2, -3, \cdots, \end{cases}$$

can be asymptotically summed for real positive x by the formula

$$\sum_{l=0}^{\infty} X_l \leqq \frac{e^{\mu\chi}}{|\,\chi^{\alpha}(2\pi\chi)^{(\mu-1)/2}\mu^{1/2}\,|},$$

where

$$\alpha \equiv \sum_i b_i - \sum_i a_i,$$

$$\chi \equiv |\,x^{1/\mu}\,|,$$

and the notation $y \leqq u$ means $\lim_{x\to\infty}(y/u) = 1$.

If we denote by $f_j(\xi)$ the series on the right side of equation (5) for $m = 0$, so that

$$y_j(t) = \alpha_{j0}\, \xi^{nj/p} f_j(\xi),$$

we see that $f_j(\xi)$ is directly amenable to the application of Stokes' formula with $x = (\xi/n)^n$. The result

---

* For the suggestion of this fundamental idea, as well as for other helpful advice, the author is indebted to Professor R. E. Langer.

† *Note on the determination of arbitrary constants*, etc., Proceedings of the Cambridge Philosophical Society, vol. 6 (1889), p. 362. The formula can also be found in Bromwich, *An Introduction to the Theory of Infinite Series*, London, 1926, Chapter XII, §112.

$$f_j(\xi) \leqq \frac{e^{|\xi|}}{(2\pi)^{\beta} n^{1/2}} \left| \left(\frac{\xi}{n} \right)^{-(nj+\nu\beta)/p} \right|$$

is then valid for x real and positive, that is, for ξ on L_{2q}, $q=0, \pm 1, \pm 2, \cdots, \pm Q/2$.

After eliminating the absolute value signs by means of the identity

$$|\xi^{\alpha}| \equiv \frac{\xi^{\alpha}}{e^{i\alpha \arg \xi}} = e^{-2\pi i \alpha q/h} \xi^{\alpha}$$

on L_{2q}, we obtain

$$y_j(t) \leqq {}_q\gamma_j \frac{e^{\omega^q \xi}}{\xi^{\nu\beta/p}} \tag{9}$$

on L_{2q}, where

$${}_q\gamma_j \equiv \frac{\kappa_j p^{nj/p} e^{(2\pi i q/n)(nj+\nu\beta)/p}}{(2\pi)^{\beta} n^{1/2-\nu\beta/p}},$$

and

$$\omega \equiv e^{-2\pi i/n}.$$

We shall now make the result (9) from Stokes' formula yield us the multipliers ${}_\tau C_{j0}$ and ${}_\tau C_{j1}$.

First, we show that in any region S_τ which abuts on a ray L_{2q}, the multiplier ${}_\tau C_{j0}$ of the dominant solution is the constant ${}_q\gamma_j$ just found. The result (9) may be written

$$y_j = {}_q\gamma_j \frac{e^{\omega^q \xi}}{\xi^{\nu\beta/p}} \{1 + \epsilon(\xi)\},$$

where $\epsilon(\xi) \to 0$ as $\xi \to \infty$ on L_{2q}. Since L_{2q} is in S_τ, the asymptotic formulas (7) are valid:

$$y_j = \sum_{k=0}^{n-1} {}_\tau C_{jk} \frac{e^{\omega_k \xi}}{\xi^{\nu\beta/p}} \left\{ 1 + \frac{E(\xi)}{\xi} \right\}$$

for $|\xi| \geqq N$; here $E(\xi)$ is a generic notation for functions of ξ, bounded for all ξ. Equating these expressions, and noting that $\omega_0 = \omega^q$ in S_τ, we obtain

$$\left| {}_\tau C_{j0} \left\{ 1 + \frac{E(\xi)}{\xi} \right\} - {}_q\gamma_j \{1 + \epsilon(\xi)\} \right| = \left| \sum_{k=1}^{n-1} {}_\tau C_{jk} e^{\omega_k \xi - \omega_0 \xi} \left\{ 1 + \frac{E(\xi)}{\xi} \right\} \right|.$$

Since on L_{2q}

$$R(\omega_k \xi - \omega_0 \xi) = h_k |\xi|, \qquad h_k < 0, \qquad k \neq 0,$$

it follows, on letting $\xi \to \infty$ on L_{2q}, that

$$ {}_{r}C_{j0} = {}_{q}\gamma_{j}. $$

We have thus determined the multipliers of the dominant solutions in all the regions S_r.

Next, we observe that for two regions S_{2q} and S_{2q-1} which abut along a ray L_{2q-1}, $-Q<2q-1<Q$, the values for ω_0 and ω_1 simply interchange as we cross L_{2q-1}:

$$ \omega_0 \text{ of } S_{2q} = \omega_1 \text{ of } S_{2q-1} = \omega^{q}, $$

$$ \omega_1 \text{ of } S_{2q} = \omega_0 \text{ of } S_{2q-1} = \omega^{q-1}. $$

We are then able to prove that the multiplier of the first sub-dominant solution on one side of L_{2q-1} is the same as the multiplier of the dominant solution on the other side of L_{2q-1}, and this latter we know. Equating

$$ y_j = \sum_{k=0}^{n-1} {}_{2q-1}C_{jk}\, {}_{2q-1}\bar{y}_k = \sum_{k=0}^{n-1} {}_{2q}C_{jk}\, {}_{2q}\bar{y}_k, $$

and utilizing the asymptotic formulas (7), which are valid on L_{2q-1} for $\{{}_{2q-1}\bar{y}_k\}$ and $\{{}_{2q}\bar{y}_k\}$, we obtain as before

$$ \left| \left\{ {}_{2q-1}C_{j0}\left[1+\frac{E(\xi)}{\xi}\right] - {}_{2q}C_{j1}\left[1+\frac{E(\xi)}{\xi}\right]\right\} \right. $$

$$ \left. + e^{\omega^{q}\xi-\omega^{q-1}\xi}\left\{ {}_{2q-1}C_{j1}\left[1+\frac{E(\xi)}{\xi}\right] - {}_{2q}C_{j0}\left[1+\frac{E(\xi)}{\xi}\right]\right\} \right| \to 0 $$

as $\xi \to \infty$ on L_{2q-1}. Now

$$ R(\omega^{q}\xi - \omega^{q-1}\xi) = 0 $$

on L_{2q-1}, and therefore

$$ \omega^{q}\xi - \omega^{q-1}\xi = ih\,|\,\xi\,|, $$

where h is real and not zero. But if

$$ A(\xi) + e^{ih|\xi|}B(\xi) \to 0 $$

as $\xi \to \infty$ and

$$ A(\xi) \to A, \qquad B(\xi) \to B, $$

then

$$ A = 0 \text{ and } B = 0. $$

Hence

$$_{2q-1}C_{j0} = {}_{2q}C_{j1} \text{ and } {}_{2q-1}C_{j1} = {}_{2q}C_{j0}.$$

The results of this section for a region S_τ are summarized in the following table:

(10)

τ	ω_0	ω_1	${}_\tau C_{j0}$	${}_\tau C_{j1}$
$2q$	ω^q	ω^{q-1}	${}_q\gamma_j$	${}_{q-1}\gamma_j$
$2q+1$	ω^q	ω^{q+1}	${}_q\gamma_j$	${}_{q+1}\gamma_j$

At this stage the asymptotic description of our special case is available. Instead of pausing for this now, however, we shall proceed to attack the general case on the basis of the results now at hand. After the treatment of the general case has been completed, results for the special case, if desired, can obviously be obtained by specialization.

CHAPTER III. A RELATED EQUATION

5. **The auxiliary variable ξ.** Although the solutions discussed in this paper are known to be single-valued in z and ρ from the general existence theorem, our asymptotic formulas involve quantities multi-valued* in z and ρ, and for this reason we introduce cuts in the z- and ρ-planes.

The function $\phi(z) \equiv z^{\nu/n}\phi_1(z)$, in general not single-valued,† can be made so with the aid of a cut in the z-plane from the origin to the boundary of R_z. The most convenient location of this cut will be specified later. Suppose tentatively some such cut, say that associated with the inequality

$$-\pi < \arg z \leqq \pi.$$

With this convention the integral

$$\Phi(z) \equiv \int_0^z \phi(z)dz \tag{11}$$

* It is the essential multi-valuedness of these quantities that gives birth to the Stokes phenomenon when $\nu>0$: Although one selects a definite branch of the multi-valued function, this eventually leads into another branch, and a change of coefficient becomes necessary to compensate for the change in the function. When $\nu=0$, the quantities are single-valued, and there is no Stokes phenomenon.

† For any given equation $\phi_1(z)$ may be chosen as an analytic function in n different ways, since only $\phi_1^n(z)$ enters the equation through hypothesis (iii). It is assumed that a definite choice is now made, and adhered to henceforth. If n divides ν, $\phi(z)$ is then single-valued, but we make the cut nevertheless.

becomes single-valued and independent of path. It is evidently of the form

$$\Phi(z) = z^{p/n}\Phi_1(z),$$

where $\Phi_1(z)$ is analytic in R_z and $\Phi_1(0) \neq 0$. The following hypothesis is indispensable:

(iv) $\Phi_1(z) \neq 0$ *in* R_z.

We now continuously deform the cut in the z-plane into the curve

$$\arg \Phi(z) = \frac{p}{n}\pi;$$

the latter is continuous with continuously turning tangent and does not cross itself. The correspondence between points Φ and z established by equation (11) can now be made one-to-one by using in the Φ-plane a Riemann surface with branch point at the origin, in the form of a sector of central angle $2\pi p/n$:

$$-\frac{p}{n}\pi < \arg \Phi \leqq \frac{p}{n}\pi.$$

The mapping of the z-plane on the Φ-plane is then conformal except at the origin, where angles are multiplied by p/n.

From the ρ-plane we delete a small circle about the origin and then cut it along a ray:*

$$|\rho| \geqq h > 0, \qquad -\pi < \arg \rho \leqq \pi. \tag{12}$$

As in Langer's work on the second-order case, we find that the key to the problem of stating in the simplest way the asymptotic formulas and their validity lies in the consideration of an auxiliary variable ξ, where we define

$$\xi \equiv \rho\Phi(z). \tag{13}$$

The Φ-plane is thus mapped on the ξ-plane by a simple magnification and rotation of amount determined by ρ.

In order to transform the t-equation into a *related equation* of the same type as the given equation, we now identify this auxiliary variable ξ with the variable ξ already introduced in §3 in connection with the t-equation. With the following definition of the number Q of (4):

$$Q \equiv 2n + \nu + \begin{cases} 0, & \nu \text{ even}, \\ 1, & \nu \text{ odd}, \end{cases}$$

* The region (12) does not include the point $\rho = \infty$.

we see that we have achieved a construction by which, in accordance with equation (11) and (13), there corresponds to any pair of points respectively in R_z and the ρ-plane a single point in the ξ-surface. At the same time we have rendered single-valued the quantities $\rho^{n/p}$, $\xi^{\nu/p}$, and $\xi^{\nu\beta/p}$, which enter into our asymptotic formulas.

6. **Transformation of t-equation into related equation.** Suppose we define $v(z, \rho)$ from the equation

$$y(t) = \psi^{\beta}(z)v(z, \rho),$$

where $y(t)$ is any solution of the t-equation, and the variables t and z are connected by equations (3) and (13). We shall prove that with proper definition of $\psi(z)$, the function $v(z, \rho)$ satisfies a related equation in R_z.

We define

$$\psi(z) \equiv \frac{\phi(z)}{\Phi^{\nu/p}(z)}. \tag{14}$$

Since arg Φ is single-valued on the Φ-surface, $\psi(z)$ is single-valued in R_z. The apparent singularity at $z=0$ is removed by the otherwise equivalent expression

$$\psi(z) = \frac{\phi_1(z)}{\Phi_1^{\nu/p}(z)},$$

which shows that $\psi(z)$ and also $\psi^{-1}(z)$ are analytic in R_z. Furthermore, since $\psi(z)$ is bounded from zero in R_z, it is possible to make a single-valued determination of $\psi^{\beta}(z)$ even when $\beta=(n-1)/2$ is not integral, and this will be analytic in R_z.

Successive differentiations* of the above equation yield

$$y^{(m)}(t) = \rho^{-mn/p}\left(\frac{p}{n}\right)^{m\nu/p}\psi^{\beta}(z)\sum_{j=0}^{m} g_{mj}(z)v^{(j)}(z, \rho), \tag{15}$$

where $g_{mj}(z)$ is a polynomial in $\psi^{-1}; \psi', \psi'', \cdots, \psi^{(n)}$ for $m \leqq n$, and is consequently analytic in R_z; in particular,

$$g_{mm}(z) = \psi^{-m}(z),$$

$$g_{m,m-1}(z) = \frac{m}{2}(n-m)\psi^{-m-1}(z)\psi'(z).$$

Substituting in the t-equation, we obtain the related equation

* The differentiations indicated by the superscripts are with respect to t on the left side of the equations and z on the right side.

$$v^{(n)}(z,\rho) + 0 + g_2(z)v^{(n-2)}(z,\rho) + \cdots + \{g_n(z) - \rho^n\phi^n(z)\}v(z,\rho) = 0, \tag{16}$$

where

$$g_j(z) \equiv \psi^n(z)g_{nj}(z)$$

is analytic in R_z.

The linear system (15), $m=0, 1, \cdots, n-1$, can be inverted,

$$v^{(m)}(z,\rho) = \psi^{-\beta}(z)\sum_{j=0}^{m}\rho^{jn/p}\left(\frac{n}{p}\right)^{jv/p}\hat{g}_{mj}(z)y^{(j)}(t), \tag{17}$$

the coefficients $\hat{g}_{mj}(z)$ being analytic in R_z, and in particular,

$$\hat{g}_{mm}(z) = \psi^m(z).$$

Those solutions of the related equation which correspond to $y(t) = {}_\tau\bar{y}_k(t)$ we shall denote by

$${}_\tau\tilde{v}_k(z,\rho) \equiv \psi^{-\beta}(z)\,{}_\tau\bar{y}_k(t).$$

Since ${}_\tau y_k^{(j)}(t)$ is analytic in

$$t = \rho^{n/p}z\left\{\frac{p}{n}\,\Phi_1(z)\right\}^{n/p},$$

it follows from (17) that ${}_\tau\tilde{v}_k^{(m)}(z,\rho)$ is an analytic function of z in R_z and of ρ in the region (12).

In the next chapter we shall need the asymptotic formulas

$${}_\tau\tilde{v}_k^{(m)}(z,\rho) = \rho^{mn/p}\psi^{m-\beta}(z)\left(\frac{n}{p}\right)^{mv/p}\left\{{}_\tau\bar{y}_k^{(m)}(t) + \frac{E(z,\rho)}{\rho^{n/p}}\right\} \tag{18}$$

for $|\xi| < N$, and

$${}_\tau\tilde{v}_k^{(m)}(z,\rho) = \rho^{mn/p}\omega_k^m\psi^{m-\beta}(z)\frac{e^{\omega_k\xi}}{\xi^{v(\beta-m)/p}}\left\{1 + \frac{E(\xi)}{\xi} + \frac{E(z,\rho)}{\rho^{n/p}}\right\} \tag{19}$$

for ξ in S_τ and $|\xi| \geqq N$. Here as elsewhere in this paper the symbol E denotes a function of the indicated variables bounded for all values of those variables. The first of the formulas follows immediately from equation (17) and the fact that ${}_\tau y_k^{(j)}(t)$, being a linear combination of $\sigma_{\tau j}(\xi)$ (see (5) et seq.) with constant coefficients, is bounded for $|\xi| < N$. The second follows from (17) together with the asymptotic expressions (7).

CHAPTER IV. COMPARISON OF SOLUTIONS

7. An equivalent integral equation. In this chapter we shall compare certain solutions of the given and related equations by an appropriate modification of a method of Birkhoff.*

* *On the asymptotic character*, etc., loc. cit., especially p. 226 ff.

Since the index τ is fixed for the course of the chapter, we shall drop it from the symbols $\{{}_\tau v_k^{(m)}(z, \rho)\}$. If $\{w_j(z, \rho)\}$ are defined from the equations*

$$\sum_{j=0}^{n-1} \tilde{v}_j^{(m)}(z, \rho) w_j(z, \rho) = \delta_{m,n-1} \qquad (m = 0, 1, \cdots, n-1),$$

it is easily seen that they are analytic functions of z in R_z and of ρ in (12). An integral equation equivalent to the given equation (1) is

$$(20) \qquad u(z, \rho) = \sum_{j=0}^{n-1} c_j \tilde{v}_j(z, \rho) + \int_0^z \left\{ \sum_{j=0}^{n-1} \tilde{v}_j(z, \rho) w_j(\zeta, \rho) \right\} L[u(\zeta, \rho)] d\zeta,$$

where

$$L[u(\zeta, \rho)] \equiv \sum_{j=0}^{n-2} \{ g_{n-j}(\zeta) - p_{n-j}(\zeta) \} u^{(j)}(\zeta, \rho),$$

and $\{c_j\}$ is a set of arbitrary constants (i.e., functions of ρ alone). If we contemplate only solutions of (20) analytic in R_z, the path of integration is evidently immaterial as long as it lies in R_z.

If the set $\{c_j\}$ is fixed, the existence and uniqueness of the solution of (20) which is analytic in z over R_z follow from the fundamental existence theorem applied to (1); for the solution of (20) is identical with that solution of (1) which obeys the boundary conditions

$$(21) \qquad u^{(m)}(0, \rho) = \sum_{j=0}^{n-1} c_j \tilde{v}_j^{(m)}(0, \rho) \quad (m = 0, 1, \cdots, n-1).$$

This solution and its derivatives are analytic in ρ if the constants $\{c_j\}$ are.

For the purpose of showing that to each solution $\tilde{v}_k(z, \rho)$ of the related equation there corresponds a solution $\tilde{u}_k(z, \rho)$ of the given equation of a certain type, the integral in equation (20) is not yet in tractable form. A suitable form of the integral equation for a definite k is

$$(22) \qquad \begin{aligned} u(z, \rho) = \sum_{j=0}^{n-1} c_j' \tilde{v}_j(z, \rho) &+ \int_Z^z \left\{ \sum_{j=0}^{k} \tilde{v}_j(z, \rho) w_j(\zeta, \rho) \right\} L[u(\zeta, \rho)] d\zeta \\ &+ \int_0^z \left\{ \sum_{j=k+1}^{n-1} \tilde{v}_j(z, \rho) w_j(\zeta, \rho) \right\} L[u(\zeta, \rho)] d\zeta, \end{aligned}$$

where Z is a fixed point of R_z; and this results from the following transformation of the constants:

---

* δ_{ij} is the "Kronecker δ."

$$c_j = c_j' - \int_0^z w_j(\zeta, \rho)L[u(\zeta, \rho)]d\zeta \qquad (j = 0, 1, \cdots, k),$$

$$c_j = c_j' \qquad (j = k+1, \cdots, n-1).$$

*The solution** $\tilde{u}_k(z, \rho)$ *is now defined to be the solution of* (22) *for* $\{c_j' = \delta_{jk}\}$. One would like to establish, for certain "admitted values" of arg ρ and $|\rho| > \rho_1$, the existence and uniqueness of this solution, together with its analyticity in ρ, after which it is easy to derive its asymptotic form. It is prerequisite, however, to inquire into the nature of suitable curves of integration and the existence of certain bounds. It is therefore convenient to conclude this section with a lemma to whose proof we can return after developing the necessary machinery.

From (21) we see that the solution of (20) determined by the set $\{c_j\}$ is

$$u(z, \rho) = \sum_{j=0}^{n-1} c_j U_j(z, \rho),$$

where $\{U_j(z, \rho)\}$ are the solutions of (1) with the boundary values

$$U_j^{(m)}(0, \rho) = \tilde{v}_j^{(m)}(0, \rho) \qquad (m = 0, 1, \cdots, n-1).$$

Hence

$$c_j' = \sum_{l=0}^{n-1} \gamma_{jl}(\rho)c_l. \tag{23}$$

The coefficients $\gamma_{jl}(\rho)$ of this transformation are analytic in ρ, since the boundary values $\{\tilde{v}_j^{(m)}(0, \rho)\}$, and therefore the solutions $\{U_j(z, \rho)\}$, and, as we have seen, the functions $\{w_j(z, \rho)\}$, all have this property.

If the following lemma is true, then, for admitted values of arg ρ and $|\rho| > \rho_1$, a unique set $\{c_j\}$ analytic in ρ will correspond to the set $\{c_j' = \delta_{jk}\}$, and consequently by the statements associated with equation (21), we shall have established the existence, uniqueness, and analyticity of $\tilde{u}_k(z, \rho)$ for such ρ.

LEMMA. *There exists a number* ρ_1 *such that, for* $|\rho| > \rho_1$ *and admitted values of* arg ρ, *the determinant* $|(\gamma_{jl}(\rho))|$ *of the system* (23) *can have no zero.*

8. **Curves** Γ, **regions** r, **and configurations** C. The quantity arg ρ now assumes critical significance, and it will simplify the notation to denote it by a single symbol λ,

$$\lambda \equiv \arg \rho.$$

* This solution is evidently a function of Z and r; we shall, however, drop these prescripts from the more precise notation ${}_{Z,r}\tilde{u}_k(z, \rho)$ until we reach the existence theorem at the end of the chapter.

By a rotation about the origin through an angle $-\lambda$, the complex of rays L_τ and sectors S_τ of the ξ-surface is mapped back on the Φ-surface, where we designate it by (cf. §3)

$L_\tau(\lambda)$: $$\arg \Phi = \theta_\tau - \lambda \equiv \theta_\tau(\lambda),$$

and

$S_\tau(\lambda)$: $$\theta_{\tau-1}(\lambda) \leqq \arg \Phi \leqq \theta_\tau(\lambda).$$

Suppose we call the map of R_z on the Φ-surface R_Φ. With such parts of the complex of rays $L_\tau(\lambda)$ and regions $S_\tau(\lambda)$ as fall outside R_Φ, we shall not be concerned.

Let Z be a fixed point of R_z, not the origin. The corresponding point $\Phi(Z)$ in R_Φ then lies in some region $S_\tau(\lambda)$, where the subscript τ is a step function of λ. The following definitions are framed relative to some definite value of τ.

A curve $\Gamma_{Z,\tau}(\lambda)$ is defined to have the following properties:

[1] *It is an ordinary curve connecting* 0 *and* Z, *lying in the cut region* R_z.

A curve is here called *ordinary* if it is continuous and has a unique tangent everywhere except possibly at a finite number of points where there exist two tangents, one for each sense of approach. A positive sense of traversal of the curve can be established; we take this to be from 0 to Z. Now consider the map of $\Gamma_{Z,\tau}(\lambda)$ on R_Φ. Let t be a unit vector directed along the tangent in the positive sense:

[2] *At every point* on the curve*

$$\theta_{\tau-1}(\lambda) \leqq \arg \mathrm{t} \leqq \theta_\tau(\lambda)$$

for some determination of arg t.

In order to see the significance of property [2], suppose z_α and z_β are two points on a Γ-curve. To signify that the curve is traversed in the positive sense in passing from z_α to z_β we shall write

$$z_\alpha \overset{\frown}{<} z_\beta.$$

Let

$$\xi_j \equiv \rho\Phi(z_j) \qquad (j = \alpha, \beta).$$

The property [2] insures that if $z_\alpha \overset{\frown}{\leqq} z_\beta$,

$$R[\omega_0(\xi_\beta - \xi_\alpha)] \geqq R[\omega_1(\xi_\beta - \xi_\alpha)] \geqq \cdots \geqq R[\omega_{n-1}(\xi_\beta - \xi_\alpha)]. \tag{24}$$

* At an exceptional point both values of t must satisfy [2].

For $n=2$, the property [2] can be replaced* by the less stringent

$$[2'] \qquad \begin{aligned} &\theta_{\tau-1}(\lambda) \leqq \arg t \leqq \theta_{\tau+1}(\lambda), && \tau \text{ even}, \\ &\theta_{\tau-2}(\lambda) \leqq \arg t \leqq \theta_{\tau}(\lambda), && \tau \text{ odd}; \end{aligned}$$

together with

[3] *The map in R_Φ of $\Gamma_{Z,\tau}(\lambda)$ lies entirely in $S_\tau(\lambda)$;*

and

[4] *There exists a number G, independent of Z, τ, λ, such that the length of all curves $\Gamma_{Z,\tau}(\lambda)$ is $<G$.*

The properties [3] and [4] are implied by [2].

A region $r_{Z,\tau}(\lambda)$ in R_z is defined to be a simply connected region containing 0 and Z, every point of which lies on some curve $\Gamma_{Z,\tau}(\lambda)$.

Although $L_\tau(\lambda)$ and $S_\tau(\lambda)$ are originally constructed on R_Φ, and $\Gamma_{Z,\tau}(\lambda)$ and $r_{Z,\tau}(\lambda)$ on R_z, no confusion will result from designating the curve or region and its map on the other surface by the same symbol.

Evidently $r_{Z,\tau}(\lambda)$ lies completely in $S_\tau(\lambda)$.

The method of comparison employed is valid for a region of the type $r_{Z,\tau}(\lambda)$. It is hence important to secure that, for each λ under consideration, every point of R_z lies in some $r_{Z,\tau}(\lambda)$. Since it may be possible to accomplish this for the desired R_z only for a range of λ more restricted than (12), we shall refer to *admitted* values of λ. Necessary and sufficient is hypothesis

(v) *For each admitted value of $\lambda \equiv \arg \rho$, every point $z \neq 0$ in R_z can be connected to the origin by a curve $\Gamma_{z,\tau}(\lambda)$.*

The condition is necessary; for, if z lies in some $r_{Z,\tau}(\lambda)$, it lies on some curve $\Gamma_{Z,\tau}(\lambda)$. The portion of the curve between 0 and z will serve as $\Gamma_{z,\tau}(\lambda)$. On the other hand, suppose z lies on $\Gamma_{z,\tau}(\lambda)$. Take $Z=z$; then z lies in an $r_{Z,\tau}(\lambda)$ consisting at least of this curve. Thus, the condition is also sufficient.

A configuration $C_{Z,\tau}$ consists, for fixed Z, τ, of all pairs of values (z, ρ) such that for each admitted $\lambda \equiv \arg \rho$ the corresponding z lies in $r_{Z,\tau}(\lambda)$.

To clarify the nature of a configuration $C_{Z,\tau}$, suppose λ_0 is an admitted value of λ, and Z any point of R_z not 0, henceforth held fixed. By the above hypothesis $\Phi(Z)$ lies on some curve $\Gamma_{Z,\tau}(\lambda_0)$. If $n>2$ the region $r_{Z,\tau}(\lambda_0)$ as seen on R_Φ will then consist at most of the parallelogram with vertices at 0 and $\Phi(Z)$ and a pair of sides along $L_{\tau-1}(\lambda_0)$, $L_\tau(\lambda_0)$. This maximum extent of

* The anomaly of the second-order case is related to the following simple geometrical problem: A circle and a straight line lie in the same plane. The circle is divided into n equal parts by the points $\bar{\omega}_0, \bar{\omega}_1, \cdots, \bar{\omega}_{n-1}$. These are projected orthogonally into the points $\alpha_0, \alpha_1, \cdots, \alpha_{n-1}$ on the line. What is the maximum range of rotation of the circle about its center under the condition that no two points α_j, α_k are allowed to cross each other? The answer is π/n for $n>2$, but $2\pi/n$ for $n=2$.

$r_{Z,\tau}(\lambda_0)$ will be possible if and only if the boundary of R_Φ does not cut across the parallelogram. The minimum extent of $r_{Z,\tau}(\lambda_0)$ will be the curve $\Gamma_{Z,\tau}(\lambda_0)$; in any case it lies entirely inside the closed region bounded by the parallelogram. All the pairs (z, ρ) for which z lies in $r_{Z,\tau}(\lambda_0)$ and $\arg \rho = \lambda_0$ will form a subset $C_{Z,\tau}(\lambda_0)$ of $C_{Z,\tau}$. If we now permit λ to vary from λ_0, the bounding parallelogram changes shape, the vertices 0, $\Phi(Z)$ remaining fixed, as well as the angles, but the sides rotating with the complex of rays $L_q(\lambda)$. There will evidently exist an interval λ_1, $\lambda_1+\pi/n$ in which the parallelogram varies between limiting positions in the form of the line segment from 0 to $\Phi(Z)$, and beyond which λ can not vary without changing the τ of our definitions. For each admitted λ in the interval, the above discussion for $\lambda=\lambda_0$ is valid and a subset $C_{Z,\tau}(\lambda)$ of $C_{Z,\tau}$ is generated. The sum of the subsets $C_{Z,\tau}(\lambda)$ for all admitted λ in the interval is identical with $C_{Z,\tau}$.

The bounding parallelogram maps conformally on R_z except at the origin where the angle π/n between the sides maps into an angle π/p between the corresponding curvilinear segments.

Because of the rotational symmetry about the origin, of the definitions with respect to integral multiples of the angle π/n, it follows that if λ_0 is an admitted value, all values $\lambda \equiv \lambda_0 \pmod{\pi/n}$ in the interval (12) are also admissible. In particular, if the whole interval λ_1, $\lambda_1+\pi/n$ is admitted, all of (12) is admissible. In this case R_Φ is characterized by the property that no ray from the origin re-enters R_Φ after once leaving it. This follows from the nature of the limiting positions of the bounding parallelograms.

If we define the curves Γ from the property [2] in the case $n=2$, the above discussion applies. If, however, we use [2'], [3], [4], the shape of the regions $r_{Z,\tau}(\lambda)$ is much more general. The implications of [1], [3], [4] are evident; [2'] implies the following characteristic: A region $r_{Z,\tau}(\lambda)$ has in common with any parallel to $L_1(\lambda)$ at most one segment (possibly a point). The modulus of congruence of the admitted values of λ then becomes π instead of π/n. However, the shape of an R_Φ valid for all λ in (12) is of the same type as above.

9. **Calculation of bounds.** Because the solutions are of two distinctly different types according as $|\xi| < N$ or $|\xi| \geqq N$, we find it useful to define a function s as follows:

$$s \equiv \begin{cases} 0, & |\xi| < N, \\ 1, & |\xi| \geqq N. \end{cases}$$

Write

$$\tilde{v}_j^{(m)}(z, \rho) = \frac{\rho^{mn/p} e^{\omega_j \xi}}{\xi^{sv(\beta-m)/p}} \hat{v}_{jm}(z, \rho) \quad (j, m = 0, 1, \cdots, n-1). \tag{25}$$

From the formulas (18) and (19) we have for ρ in (12) and ξ in S_τ,

$$|\hat{v}_{jm}(z, \rho)| < M_1.$$

Again, if we write

$$w_j(z, \rho) = \frac{\hat{w}_j(z, \rho)}{\rho^{2\beta n/p}\xi^{s\nu\beta/p}e^{\omega_j\xi}}, \tag{26}$$

we see that the $\{\hat{w}_j(z, \rho)\}$ are determined from the equations

$$\sum_{j=0}^{n-1} \hat{v}_{jm}(z, \rho)\hat{w}_j(z, \rho) = \delta_{m,n-1} \qquad (m = 0, 1, \cdots, n-1).$$

By solving (25) for the element $\hat{v}_{jm}(z, \rho)$ and utilizing the formula (19), the determinant $\Delta(z, \rho)$ of this system is obtained in the form

$$\Delta(z, \rho) = |(\omega_j^m)| + \frac{E(\xi)}{\xi} + \frac{E(z, \rho)}{\rho^{n/p}}$$

for $|\xi| \geqq N$ and ξ in S_τ. The well known matrix (ω_j^m), $j, m = 0, 1, \cdots, n-1$, cannot have a zero determinant, and hence, after possibly necessary readjustment of the constant N to a larger value, the determinant $\Delta(z, \rho)$ is bounded from zero for $|\rho| >$ some ρ_2, ξ in S_τ, and $|\xi| \geqq N$. A similar result can be obtained for $|\xi| < N$, since, by (18),

$$\Delta(z, \rho) = \left(\frac{n}{p}\right)^{\nu n\beta/p} W[{}_\tau\bar{y}_j(t)] + \frac{E(z, \rho)}{\rho^{n/p}}.$$

Here W is the Wronskian of the solutions $\{{}_\tau\bar{y}_j(t)\}$, and since these satisfy an equation in normal form, $W =$ constant, and since they are independent, $W \neq 0$. We thus conclude for $|\rho| > \rho_3$, and ξ in S_τ,

$$|\hat{v}_{jm}(z, \rho)| < M, \qquad |\hat{w}_j(z, \rho)| < M. \tag{27}$$

Finally, if we write

$$\begin{aligned} \sum_{j=0}^{k} \bar{v}_j^{(m)}(z, \rho)w_j(\zeta, \rho) &\equiv \frac{\rho^{n(m-2\beta)/p}e^{\omega_k(\xi-\bar{\xi})}}{\xi^{s\nu(\beta-m)/p}\bar{\xi}^{\bar{s}\nu(\beta-m)/p}} A_{km}(z, \zeta, \rho) \\ &\equiv F_{km}(z, \zeta, \rho)A_{km}(z, \zeta, \rho), \end{aligned} \tag{28}$$

$$\sum_{j=k+1}^{n-1} \bar{v}_j^{(m)}(z, \rho)w_j(\zeta, \rho) \equiv F_{km}(z, \zeta, \rho)B_{km}(z, \zeta, \rho),$$

where it is now necessary to distinguish between $\xi \equiv \rho\Phi(z)$ and $\bar{\xi} \equiv \rho\Phi(\zeta)$, likewise $s \equiv f(\xi)$ and $\bar{s} \equiv f(\bar{\xi})$, we can prove that the functions A_{km} and B_{km} are bounded in the following manner for $|\rho| > \rho_3$ and ξ, $\bar{\xi}$ in S_τ:

$$(29)\qquad \begin{aligned} &|A_{km}(z,\zeta,\rho)| < T, \qquad 0 \leqq z \leqq \zeta \leqq Z,\\ &|B_{km}(z,\zeta,\rho)| < T, \qquad 0 \leqq \zeta \leqq z \leqq Z, \end{aligned}$$

where T is independent of Z. From (25), (26), (27),

$$|\bar{v}_j^{(m)}(z,\rho)w_j(\zeta,\rho)| < |F_{jm}(z,\zeta,\rho)|\, M^2.$$

But from (24) we know

$$|F_{jm}(z,\zeta,\rho)| \leqq |F_{km}(z,\zeta,\rho)|$$

with the restriction

$$\begin{aligned} &0 \leqq z \leqq \zeta \leqq Z && (j = 0, 1, \cdots, k),\\ &0 \leqq \zeta \leqq z \leqq Z && (j = k+1, \cdots, n-1). \end{aligned}$$

Combining these inequalities with (28), we have (29).

We are now equipped with the necessary bounds to prove our lemma. It is convenient, however, first to write the equation (22) and those resulting from differentiation of (22) in a form exhibiting the functions $\{\hat{u}_m(z,\rho)\}$ which are defined from

$$(30)\qquad u^{(m)}(z,\rho) \equiv \frac{\rho^{mn/p}e^{\omega_k\xi}}{\xi^{\alpha\nu(\beta-m)/p}}\hat{u}_m(z,\rho);$$

namely, with $m=0, 1, \cdots, n-1$,

$$(31)\qquad \begin{aligned} \hat{u}_m(z,\rho) = \sum_{j=0}^{n-1} c_j' e^{(\omega_j-\omega_k)\xi}\hat{v}_{jm}(z,\rho) &+ \int_Z^z A_{km}(z,\zeta,\rho)\hat{L}[u(\zeta,\rho)]\,d\zeta\\ &+ \int_0^z B_{km}(z,\zeta,\rho)\hat{L}[u(\zeta,\rho)]d\zeta, \end{aligned}$$

where

$$\hat{L}[u(\zeta,\rho)] \equiv \sum_{j=0}^{n-2} \frac{\{g_{n-j}(\zeta) - p_{n-j}(\zeta)\}\hat{u}_j(\zeta,\rho)}{\rho^{n(2\beta-j)/p}\xi^{\alpha\nu(2\beta-j)/p}}.$$

10. **Proof of lemma; existence theorem.** We now specify the path of integration to be a curve $\Gamma_{Z,\tau}(\lambda)$ whenever λ is an admitted value. Suppose for some ρ with admitted λ the determinant $|(\gamma_{jl}(\rho))|$ has a zero. Then with $\{c_j'\}$ all zero, a set $\{c_j\}$ not all zero can be found to satisfy (23). That is, for this ρ there exists a solution $u(z,\rho)\not\equiv 0$ of the equation (22) with $\{c_j'\}$ all zero.

Let $U(\rho)\neq 0$ be the maximum of $|\hat{u}_m(z,\rho)|$, $m=0, 1, \cdots, n-1$, in the region $r_{Z,\tau}(\lambda)$. Since the coefficients $g_l(z)$ and $p_l(z)$ are analytic in the closed region R_z,

$$| g_l(z) - p_l(z) | < D \qquad (l = 2, 3, \cdots, n);$$

and hence from (31),*

$$| \hat{u}_m(z, \rho) | < (n-1) \cdot T \cdot D \cdot G \cdot \frac{U(\rho)}{| \rho |^{n/p}}.$$

But for some $m = m_1(\rho)$ and some $z = z_1(\rho)$,

$$\frac{n-1}{n} U(\rho) < | \hat{u}_m(z, \rho) |.$$

Combining the last two inequalities, we see our lemma is verified with

$$\rho_1 \equiv \{ n \cdot T \cdot D \cdot G \}^{p/n}.$$

With the machinery constructed to prove the lemma, we can easily derive the asymptotic forms of $\{ \tilde{u}_k^{(m)}(z, \rho) \}$ for (z, ρ) in the configuration $C_{Z,\tau}$. If the constants $\{ c_j' \}$ of (22) are specified to be $\{ \delta_{jk} \}$, the solution $u(z, \rho)$ becomes identical with $\tilde{u}_k(z, \rho)$,

$$u^{(m)}(z, \rho) \equiv \tilde{u}_k^{(m)}(z, \rho).$$

With a definition of $U(\rho)$ analogous to the previous, we obtain from (31), for admitted λ,

$$| \hat{u}_m(z, \rho) | < M + (n-1) \cdot T \cdot D \cdot G \cdot \frac{U(\rho)}{| \rho |^{n/p}},$$

and, as before, the left side may be replaced for some z, m by $(n-1)U(\rho)/n$. Hence

$$U(\rho) < \frac{2n}{n-1} M$$

for

$$| \rho | > \rho_0 \equiv 2^{p/n} \rho_1.$$

We conclude now from (31) that, for $| \rho | > \rho_0$,

$$| \hat{u}_m(z, \rho) - \hat{v}_{km}(z, \rho) | < \frac{2n \cdot M \cdot T \cdot D \cdot G}{| \rho |^{n/p}}.$$

On noting that this bound and ρ_0 are independent of Z, we are led to

---

* Note $2\beta - j \geqq 1$ if $j \leqq n-2$; also recall the property [4] that G is a bound for the length of all Γ-curves.

THEOREM 1. *Corresponding to any configuration $C_{Z,\tau}$ there exist n solutions $\{{}_{Z,\tau}\bar{u}_k(z, \rho)\}$ of the given equation, which, for (z, ρ) in $C_{Z,\tau}$ and $|\rho| > \rho_0$, are analytic in ρ and satisfy the relations*

$$(32) \quad {}_{Z,\tau}\bar{u}_k^{(m)}(z, \rho) = {}_{\tau}\bar{v}_k^{(m)}(z, \rho) + \rho^{(m-1)n/p}E(z, \rho) \quad (k, m = 0, 1, \cdots, n-1)$$

when $|\xi| < N$, and

$$(33) \quad {}_{Z,\tau}\bar{u}_k^{(m)}(z, \rho) = \rho^{mn/p}\omega_k^m\psi^{m-\beta}(z)\frac{e^{\omega_k\xi}}{\xi^{\nu(\beta-m)/p}}\left\{1 + \frac{E(\xi)}{\xi} + \frac{E(z, \rho)}{\rho^{n/p}}\right\}$$

when $|\xi| \geqq N$. Here the E-functions are bounded uniformly for all configurations $C_{Z,\tau}$.

The independence of these solutions is evident from their forms (33).

CHAPTER V. BOUNDARY VALUE PROBLEM AT THE ZERO

11. **Forms for the solutions principal at the zero.** The solutions $\{u_j(z, \rho)\}$ of the given equation, analytic in z and ρ, which satisfy the boundary conditions

$$u_j^{(m)}(0, \rho) = \delta_{jm} \qquad (j, m = 0, 1, \cdots, n-1),$$

constitute the set we call principal at the origin. These must be linear combinations of the solutions described in Theorem 1,

$$(34) \qquad u_j(z, \rho) = \sum_{k=0}^{n-1} {}_{Z,\tau}\beta_{jk}(\rho)\, {}_{Z,\tau}\bar{u}_k(z, \rho).$$

In order to derive the form of ${}_{Z,\tau}\beta_{jk}(\rho)$ for admitted $\lambda \equiv \arg \rho$ and $|\rho| > \rho_0$, let us hold Z, τ fixed until further notice. Meanwhile, we shall drop these subscripts wherever they occur.

For $|\xi| < N$, (z, ρ) in the configuration C, and $|\rho| > \rho_0$, we have from (32)

$$u_j^{(m)}(z, \rho) = f_m(z, \rho)\sum_{k=0}^{n-1}\beta_{jk}(\rho)\left\{y_k^{(m)}(t) + \frac{E(z, \rho)}{\rho^{n/p}}\right\},$$

where

$$f_m(z, \rho) \equiv \rho^{mn/p}\psi^{m-\beta}(z)\left(\frac{n}{p}\right)^{m\nu/p}.$$

Since $\{y_j\}$ and $\{\bar{y}_k\}$ are each sets of n independent solutions, the system (8) can be inverted,

$$\bar{y}_k = \sum_{l=0}^{n-1} \mathfrak{c}_{kl}y_l \qquad (k = 0, 1, \cdots, n-1),$$

where

$$\sum_{k=0}^{n-1} c_{jk}\hat{c}_{kl} = \delta_{jl}.$$

Consequently

$$(35) \qquad u_j^{(m)}(z, \rho) = f_m(z, \rho)\sum_{k=0}^{n-1} \beta_{jk}(\rho)\left\{\left[\sum_{l=0}^{n-1} \hat{c}_{kl}y_l^{(m)}\right] + \frac{E(z, \rho)}{\rho^{n/p}}\right\}.$$

Now let $z=0$; then $t=0$, $y_l^{(m)}=\delta_{lm}$, and thus

$$\delta_{jm} = f_m(0, \rho)\sum_{k=0}^{n-1} \beta_{jk}(\rho)\left\{\hat{c}_{km} + \frac{E(\rho)}{\rho^{n/p}}\right\}.$$

This may be written

$$\delta_{jm} = \sum_{k=0}^{n-1}\{f_j(0, \rho)\beta_{jk}(\rho)\}\left\{\hat{c}_{km} + \frac{E(\rho)}{\rho^{n/p}}\right\}.$$

Compare with this the system

$$\delta_{jm} = \sum_{k=0}^{n-1} \chi_{jk}\hat{c}_{km} \qquad (m = 0, 1, \cdots n-1),$$

whose solutions are evidently

$$\chi_{jk} = c_{jk}.$$

It follows that a system

$$\delta_{jm} = \sum_{k=0}^{n-1} \chi_{jk}(\rho)\left\{\hat{c}_{km} + \frac{E(\rho)}{\rho^{n/p}}\right\}$$

has solutions

$$\chi_{jk} = c_{jk} + \frac{E(\rho)}{\rho^{n/p}};$$

and hence

$$(36) \qquad \beta_{jk}(\rho) = \frac{1}{f_j(0, \rho)}\left\{c_{jk} + \frac{E(\rho)}{\rho^{n/p}}\right\}.$$

Substitution of this value in (35) and interchange of the order of summation yields the equation

$$u_j^{(m)}(z, \rho) = \frac{f_m(z, \rho)}{f_j(0, \rho)}\left\{\left[\sum_{l=0}^{n-1} y_l^{(m)}\sum_{k=0}^{n-1} c_{jk}\hat{c}_{kl}\right] + \frac{E(z, \rho)}{\rho^{n/p}}\right\},$$

or

$$u_j^{(m)}(z, \rho) = \frac{f_m(z, \rho)}{f_j(0, \rho)}\left\{y_j^{(m)}(t) + \frac{E(z, \rho)}{\rho^{n/p}}\right\}.$$

If we now resume the subscripts Z, τ, we find that in the last equation they apply only to the E-function. But we were careful to note that the E-functions of Theorem 1 were uniformly bounded for all Z, τ. Consequently the present E-function inherits the same property, that is, it is bounded uniformly for all (z, ρ) for which z is in R_z, $|\rho| > \rho_0$, λ is an admitted value, and $|\xi| < N$. On this result we base

THEOREM 2. *For admitted values of* $\arg \rho$, $|\rho| > \rho_0$, *and* z *in* R_z, *the solutions of the given equation principal at the origin,* $\{u_j^{(m)}(0, \rho) = \delta_{jm},\ j, m = 0, 1, \cdots, n-1\}$, *have the form*

$$u_j^{(m)}(z, \rho) = \rho^{(m-j)n/p}\kappa_j p^{jn^{m-j}}\frac{\psi^{m-\beta}(z)}{\psi^{j-\beta}(0)}\left\{\sigma_{jm}(\xi) + \frac{E(z, \rho)}{\rho^{n/p}}\right\} \tag{37}$$

when $|\xi| < N$. *The series* $\sigma_{jm}(\xi)$ *and constants* κ_j *are defined by equations* (5) *et seq.*

In unabbreviated notation, equation (36) becomes

$${}_{Z,\tau}\beta_{jk}(\rho) = \frac{\rho^{-jn/p}\left(\frac{p}{n}\right)^{jv/p}}{\psi^{j-\beta}(0)}\left\{{}_{\tau}c_{jk} + \frac{E(\rho)}{\rho^{n/p}}\right\},$$

where the E-function is uniformly bounded for all Z, τ. Taking this and the forms (33) into (34), we obtain

$$u_j^{(m)}(z, \rho) = \rho^{(m-j)n/p}\left(\frac{p}{n}\right)^{jv/p}\frac{\psi^{m-\beta}(z)}{\psi^{j-\beta}(0)}\,\frac{{}_{\tau}H_{jm}(z, \rho)}{\xi^{v(\beta-m)/p}},$$

where

$${}_{\tau}H_{jm}(z, \rho) \equiv \sum_{k=0}^{n-1}\omega_k^m e^{\omega_k\xi}\left\{{}_{\tau}c_{jk} + \frac{E(\xi)}{\xi} + \frac{E(z, \rho)}{\rho^{n/p}}\right\},$$

and, for admitted λ and $|\rho| > \rho_0$, the E-functions are bounded uniformly when ξ is in S_τ and $|\xi| \geqq N$.

We recall that a region T_q is bisected by the ray L_{2q} (see §3), but that in both halves the number ω_0 of the dominant solution has the same value ω^q (see (10)). In T_q

$$R(\omega_k\xi - \omega_0\xi) \leqq h|\xi|, \qquad h < 0, \qquad k \neq 0;$$

also, the multiplier of the dominant solution is the same in both halves of T_q,

namely, ${}_q\gamma_j \neq 0$; hence with $\tau = 2q$ or $2q+1$, depending on which half of T_q contains ξ,

$$ {}_\tau H_{jm}(z, \rho) = {}_q\gamma_j \omega^{qm} e^{\omega^q \xi} \left\{ 1 + \frac{E(\xi)}{\xi} + \frac{E(z, \rho)}{\rho^{n/p}} \right\}. $$

By a similar method we can treat the case when ξ is in the sector $D_{q,q+1}$ between T_q and T_{q+1}; the results are contained in

THEOREM 3. *When* $|\xi| \geqq N$, *and for admitted values of* arg ρ, $|\rho| > \rho_0$, *and* z *in* R_z, *the principal solutions of the given equation are of the form*

$$ (38) \qquad u_j^{(m)}(z, \rho) = {}_q\Psi_{jm}(z, \rho) \left\{ 1 + \frac{E(\xi)}{\xi} + \frac{E(z, \rho)}{\rho^{n/p}} \right\} $$

for ξ *in* T_q, *and*

$$ (39) \qquad u_j^{(m)}(z, \rho) = {}_q\Psi_{jm}(z, \rho) \left\{ 1 + \frac{E_1(\xi)}{\xi} + \frac{E_1(z, \rho)}{\rho^{n/p}} \right\} + {}_{q+1}\Psi_{jm}(z, \rho) \left\{ 1 + \frac{E_2(\xi)}{\xi} + \frac{E_2(z, \rho)}{\rho^{n/p}} \right\} $$

for ξ *in* $D_{q,q+1}$, *where*

$$ (40) \qquad {}_q\Psi_{jm}(z, \rho) \equiv \frac{\rho^{(m-j)n/p} \kappa_j p^j e^{(2\pi i q/n)[(nj+\nu\beta)/p - m]}}{n^{1/2+\nu(j-\beta)/p}(2\pi)^\beta} \cdot \frac{\psi^{m-\beta}(z)}{\psi^{j-\beta}(0)} \cdot \frac{e^{\omega^q \xi}}{\xi^{\nu(\beta-m)/p}}. $$

12. **Discussion of results.** Two sectors T_q and T_{q+n} lie in the same position but in adjacent sheets of the ξ-surface. The analytic expressions given by Theorem 3 for $u_j^{(m)}(z, \rho)$ differ for the two regions by a factor $e^{2\pi i \alpha_j}$, where

$$ \alpha_j \equiv \frac{nj + \nu\beta}{p} = \frac{2nj + \nu(n-1)}{2(n+\nu)}; $$

that is, the expression for ξ in T_{q+n} is $e^{2\pi i \alpha_j}$ times that for ξ in T_q. This factor may be regarded as a measure of the Stokes phenomenon.

We see that α_j is independent of the region T_q as well as the order m of the derivative. Evidently $e^{2\pi i \alpha_j}$ will be unity, in other words, there will be no Stokes phenomenon, if and only if α_j is integral. Thus, if $\nu = 0$, none of the solutions will show a Stokes phenomenon. However, if $\nu > 0$, not all the solutions can lack it, since α_j and α_{j+1} differ by n/p and hence the α_j's cannot all be integers. In particular, for $n = 2$ every principal solution manifests the

phenomenon, but for any $n>2$ there are always some (ν, j) for which it fails to appear.*

The results of Theorem 3 could have been expressed in the form†

$$u_j^{(m)}(z,\rho) = \rho^{(m-j)n/p}\left(\frac{p}{n}\right)^{j\nu/p}\frac{\psi^{m-\beta}(z)}{\psi^{j-\beta}(0)} \cdot\frac{1}{\xi^{\nu(\beta-m)/p}}\sum_{l=0}^{n-1}{}_\tau d_{jl}\omega^{lm}e^{\omega^l\xi}\left\{1+\frac{E(\xi)}{\xi}+\frac{E(z,\rho)}{\rho^{n/p}}\right\} \tag{41}$$

for $|\xi|\geqq N$, where the multiplier ${}_\tau d_{jl}$ depends on which region S_τ contains ξ. The set $\{{}_\tau d_{jl}\}$ is a permutation of the set $\{{}_\tau c_{jl}\}$, such that, if $\omega_\tau=\omega^l$ in S_τ, then ${}_\tau c_{jr}={}_\tau d_{jl}$. To see that this is entirely equivalent to Theorem 3, we need only note that if ξ is in T_q, the sum on the right of (41) may be replaced by the dominant term, the other terms being absorbed into the corresponding $E(z,\rho)$, giving formula (38). If ξ is in a D-sector, two terms must be kept as in formula (39).

Consider one of the terms of (41) for fixed l. It is dominant in the regions T_{q_l}, where $q_l\equiv l(\text{mod } n)$. The sectors T_{q_l} are all in the same position but in different sheets of the ξ-surface, and are bisected by the rays L_{2q_l}. The rays L_{τ_l}, $\tau_l\equiv 2l-n \pmod{2n}$, leave the origin in the opposite direction. With reference to the equation (41), the Stokes phenomenon could evidently be described as follows: The value of the multiplier ${}_\tau d_{jl}$ need be changed only when crossing a ray L_{τ_l}, and then by the factor $e^{2\pi i\alpha_j}$. Instead of this particular ray, it would be equally compatible with our results to employ as Stokes line any simple curve from 0 to ∞ on the ξ-surface not lying in T_{q_l} and the adjacent D-sectors. However, this location of the Stokes lines is more general than would be warranted if it were not for a certain insufficiency in our results, which we now bring to attention.

Let $u(z,\rho)$ be the solution of the given equation satisfying the boundary conditions

$$u^{(m)}(0,\rho) = h_m(\rho) \qquad (m=0,1,\cdots,n-1).$$

Then

$$u^{(m)}(z,\rho) = \sum_{j=0}^{n-1} h_j(\rho)u_j^{(m)}(z,\rho).$$

---

* If $n=2$, $\nu>0$, then $0<\alpha_j<1$, $j=0,1$.

If $n>2$, $\nu>0$, we have the following:

Case 1: n odd. For every ν at least one α_j is integral, namely, for $j=(n-1)/2$.

Case 2: n even, ν odd. No α_j can be integral, since $\alpha_j=$odd/even.

Case 3: n even, ν even. There may or may not exist integral α_j for a definite ν. However, for $\nu=n(n-2)$, α_0 and α_{n-1} are integral.

† It is understood that the discussion in this section applies only to such ρ as have admitted $\lambda\equiv\arg\rho$ and $|\rho|>\rho_0$.

Using the forms of Theorem 3, we have for $|\xi| \geqq N$ and ξ in T_q,

$$u^{(m)}(z, \rho) = {}_qA_m(z, \rho)\sum_{j=0}^{n-1} h_j(\rho)\ {}_qB_j(\rho)\left\{1 + \frac{E_j(\xi)}{\xi} + \frac{E_j(z, \rho)}{\rho^{n/p}}\right\}, \tag{42}$$

where

$${}_qB_j(\rho) \equiv \rho^{-jn/p}\ \frac{\kappa_j p^j e^{2\pi i q j/p}}{n^{pj/p}\psi^j(0)}.$$

In order that the expression (42) be explicit to a factor

$$\left\{1 + \frac{E(\xi)}{\xi} + \frac{E(z, \rho)}{\rho^{n/p}}\right\}, \tag{43}$$

the following condition on the boundary values $\{h_j(\rho)\}$ is sufficient, and, in the absence of further knowledge of the E_j-functions, also necessary: The quantities

$${}_qC_j(\rho) \equiv \frac{\rho^{-jn/p}h_j(\rho)}{\sum_{l=0}^{n-1} {}_qB_l(\rho)h_l(\rho)} \qquad (j = 0, 1, \cdots, n-1)$$

are bounded. This follows from noting that the expression (42) is explicit to a factor

$$1 + \sum_{j=0}^{n-1} {}_qC_j(\rho)\left\{\frac{\overline{E}_j(\xi)}{\xi} + \frac{\overline{E}_j(z, \rho)}{\rho^{n/p}}\right\},$$

where the $\overline{E}$'s are constant multiples of the E's. For such ρ as satisfy the conditions, and $|\xi| \geqq N$, $u^{(m)}(z, \rho)$ then has the form

$$u^{(m)}(z, \rho) = \left\{{}_qA_m(z, \rho)\sum_{j=0}^{n-1} h_j(\rho)\ {}_qB_j(\rho)\right\}\left\{1 + \frac{E(\xi)}{\xi} + \frac{E(z, \rho)}{\rho^{n/p}}\right\}$$

when ξ is in T_q, while for ξ in $D_{\bar{q},\bar{q}+1}$ the form is* the sum of two such expressions written respectively for $q = \bar{q}$ and $\bar{q}+1$. From this it is easily seen that our formulas are not explicit to a factor (43) in a region which is not zero-free for $u^{(m)}(z, \rho)$, $m = 0, 1, \cdots, n-1$.

This insufficiency of our formulas to describe the most general solution bears on the freedom of locating the Stokes lines: The only time we would need to know ${}_\tau d_{jl}$ outside a region T_{q_l} and the adjoining D-sectors would be when we have such a combination of the principal solutions,

$$u^{(m)}(z, \rho) = \sum_{j=0}^{n-1} h_j(\rho)u_j^{(m)}(z, \rho) = \sum_{k=0}^{n-1} z, {}_\tau\beta_k(\rho)\ z, {}_\tau\tilde{u}_k^{(m)}(z, \rho)$$

* Providing ${}_qC_j(\rho)$ is bounded for $q = \bar{q}$ and $\bar{q}+1$.

that the dominant solution ${}_{z,\tau}\hat{u}_0^{(m)}(z, \rho)$ fails to dominate the expression. But this can happen only in a case not covered by our formulas because of a lack of knowledge about the E-functions. If it were not for this lack, the location of the Stokes lines would have to be more restricted; thus, for the t-equation it can be shown that the Stokes lines for ${}_{\tau}d_{jl}$ must be taken to lie in those regions of the ξ-surface which are reflections of T_{q_l} in the origin.

We shall conclude with some mention of how our results specialize. Consider the case when $\nu=0$. We have already seen that there is then no Stokes phenomenon. A great simplification of the results can now be made since it is possible to assume $|\xi| \geqq N$ always, by taking $N=0$. The constant N was introduced first in §3, where it can be taken as zero, since the series

$$s_k^{(m)}(t) = \omega_k^m e^{\omega_k \xi}\left\{1 + \frac{0}{\xi} + \frac{0}{\xi^2} + \cdots\right\}$$

are now true solutions,*

$${}_{\tau}\bar{y}_k^{(m)}(t) = s_k^{(m)}(t).$$

We again encounter it in §6, where we find formula (19) with $E(\xi)\equiv 0$ is valid for all ξ. The comparison in Chapter IV can now be carried through with $N=0$, and again the formula (33) of Theorem 1 works for all ξ with $E(\xi)\equiv 0$. The method of deriving the forms of the multipliers ${}_{z,\tau}\beta_{jk}(\rho)$ in §11 is then applicable without the assumption $|\xi| < N$. We can now dispense with Theorem 2 and use Theorem 3 for all ξ with $N=0$, noting that

$$E(\xi) \equiv E_1(\xi) \equiv E_2(\xi) \equiv 0,$$

and that the formula (40) now assumes the simple form

$${}_{q}\Psi_{jm}(z, \rho) = \rho^{m-j}\,\frac{e^{2\pi i q(j-m)/n}}{n}\cdot\frac{\phi^{m-\beta}(z)}{\phi^{j-\beta}(0)}\,e^{\omega^q \xi}.$$

Finally, a word about the case when the coefficient $\phi^n(z)$ is real on the real axis, or real except for a constant imaginary factor. By the transfer of a constant factor to the parameter ρ^n, ${\phi_1}^n(z)$ may be made real and positive on the real axis. The cut in the z-plane can now be taken along the negative real axis and need not be deformed. The positive and negative real axis map respectively into the rays $\arg \Phi = 0, p/n\,\pi$. It follows that if one is interested only in real z, all values (12) of $\arg \rho$ may be admitted.

* Note $\xi \equiv t$; also $p=n$ and $\psi(z)\equiv\phi(z)$.

UNIVERSITY OF WISCONSIN,
MADISON, WIS.

DIFFERENTIAL GEOMETRY OF A SURFACE AT A PLANAR POINT*

BY

V. G. GROVE

1. Introduction

We shall say that a point O of a surface S is a *planar point* of the third order if the surface has a tangent plane at O and if every curve on the surface through O has an inflexion† at O. It is the purpose of this paper to study the surface S in the neighborhood of such a planar point from a projective point of view by making use of the osculants‡ of Bompiani of a plane section of S through an arbitrary tangent line through O.

Lane and Su§ have used Bompiani's osculants to study the plane sections of a surface at a non-parabolic point on the surface. In a sense every tangent line to a surface at a planar point is an asymptotic tangent.

Downs‖ has studied planar points of order n from a metric point of view.

In the last section we discuss the loci of certain points and lines intrinsically associated with the sections of the surface through the arbitrary tangent line as this line generates a pencil. The loci so obtained are therefore covariant loci associated with the surface and not merely with the surface in its relation to the particular tangent line.

We first find a canonical power series expansion for one non-homogeneous projective coordinate in terms of the two others. A complete geometrical description will be given of the tetrahedron of reference giving rise to this canonical expansion.

2. The power series expansion

Let the equation of the analytic surface S be

$$z = f(x, y).$$

* Presented to the Society, April 10, 1936; received by the editors October 14, 1935, and in revised form, January 11, 1936.

† Each branch of the curve of intersection of the surface with its tangent plane at a planar point of third order has at least four-point contact with its tangent line at the point.

‡ E. Bompiani, *Per lo studio proiettivo-differenziale delle singolarita*, Bollettino dell' Unione Matematica, vol. 5 (1926), p. 118. Hereafter referred to as Bompiani.

§ E. P. Lane, *Plane sections through an asymptotic tangent of a surface*, Bulletin of the American Mathematical Society, vol. 41 (1935), pp. 285–290; B. Su, *On certain quadratic cones projectively connected with a space curve and a surface*, Tôhoku Mathematical Journal, vol. 38 (1933), pp. 233–244.

‖ T. L. Downs, *Asymptotic and principal directions at a planar point of a surface*, Duke Mathematical Journal, vol. 1 (1935), pp. 316–327.

Let the system of coordinates be so chosen that the point $O(0, 0, 0)$ is the planar point, and $z=0$ the equation of the tangent plane to S at O. The Taylor's expansion of the function $f(x, y)$ in the neighborhood of $x=0$, $y=0$ will be of the form

$$z = A_3(x, y) + A_4(x, y) + A_5(x, y) + \cdots, \tag{1}$$

wherein

$$\begin{aligned} A_3(x, y) &= a_{30}x^3 + a_{21}x^2y + a_{12}xy^2 + a_{03}y^3, \\ A_4(x, y) &= a_{40}x^4 + a_{31}x^3y + a_{22}x^2y^2 + a_{13}xy^3 + a_{04}y^4, \\ A_5(x, y) &= a_{50}x^5 + a_{41}x^4y + a_{32}x^3y^2 + a_{23}x^2y^3 + a_{14}xy^4 + a_{05}y^5, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots. \end{aligned} \tag{2}$$

The tangent plane $z=0$ at O intersects the surface in a curve with a triple point at O. The triple point tangents are given by

$$A_3(x, y) = 0. \tag{3}$$

The equation $A_3(x, y)=0$ defines a cubic involution in the pencil of tangent lines to the surface at O, the triple lines being the triple point tangents to the curve of intersection of the surface S with its tangent plane at O. We shall assume that the tangent line $y=z=0$ is not one of the triple lines of the cubic involution defined by $A_3(x, y)=0$. It follows that $a_{30}\neq 0$.

We propose first to derive a canonical form of the expansion (1) in which the edge $y=z=0$ of the tetrahedron of reference giving rise to the expansion is an arbitrary tangent line to the surface S at O except for the limitation noted above.

If we make the transformation

$$x = x' + \mu y', \qquad y = y', \qquad z = z', \tag{4}$$

on equation (1), and denote the new coefficients by $\bar{a}_{30}, \bar{a}_{21}, \cdots$, we find that

$$\begin{aligned} &\bar{a}_{30} = a_{30}, \ \bar{a}_{21} = 3\mu a_{30} + a_{21}, \ \bar{a}_{12} = 3\mu^2 a_{30} + 2\mu a_{21} + a_{12}, \\ &\bar{a}_{40} = a_{40}, \ \bar{a}_{31} = 4\mu a_{40} + a_{31}, \ \bar{a}_{50} = a_{50}, \\ &\bar{a}_{41} = 5\mu a_{50} + a_{41}, \ \bar{a}_{60} = a_{60}, \cdots. \end{aligned} \tag{5}$$

Hence if we choose μ to satisfy the equation

$$3\mu a_{30} + a_{21} = 0, \tag{6}$$

we may make $\bar{a}_{21}=0$. We shall assume that this transformation has been effected on the series (1).

We may give a simple geometrical characterization of the tangent line $x=z=0$ for the case in which $a_{21}=0$. *The unique line of the involution defined*

by $A_3(x, y)=0$ corresponding to the line $y=z=0$, considered as a double line of the involution, is the line $x=z=0$ if and only if $a_{21}=0$.

The transformation

$$(7)\qquad x=\frac{x'+\alpha z'}{1+Lx'+My'},\qquad y=\frac{y'+\beta z'}{1+Lx'+My'},\qquad z=\frac{z'}{1+Lx'+My'},$$

does not change the edges $x=z=0$, $y=z=0$ of the coordinate system, and does not affect the coefficients of the form $A_3(x, y)$. If again we denote the new coefficients by $\bar{a}_{40}, \bar{a}_{31}, \cdots$, we find that

$$(8)\qquad \begin{aligned} \bar{a}_{40} &= -2La_{30}+a_{40},\quad \bar{a}_{31}=-2Ma_{30}+a_{31},\\ \bar{a}_{50} &= 3\alpha a_{30}^2-3La_{40}+a_{50},\\ \bar{a}_{41} &= 2\beta a_{30}a_{12}-3Ma_{40}+a_{41},\\ \bar{a}_{05} &= 3\beta a_{03}^2-3Ma_{04}+a_{05},\ \cdots. \end{aligned}$$

We note from (8), that, with proper choice of α, L, M, we may make $\bar{a}_{40}=\bar{a}_{31}=\bar{a}_{50}=0$. If $a_{12}\neq 0$, we may choose β in such a way that $\bar{a}_{41}=0$. If $a_{12}=0$, we may choose β so that $\bar{a}_{05}=0$.

The Hessian H of the form

$$A_3(x, y)=a_{30}x^3+a_{12}xy^2+a_{03}y^3$$

is readily shown to be

$$(9)\qquad H=4[3a_{30}a_{12}x^2+9a_{30}a_{03}xy-a_{12}^2y^2].$$

Hence *the tangent line $y=z=0$ ($x=z=0$) is one of the Hessian or neutral lines of the cubic involution determined by $A_3(x, y)=0$ if and only if $a_{12}=0$.*

We shall say that $y=z=0$ is a *Hessian tangent* if $a_{12}=0$.

Let us suppose first that $y=z=0$ is not a Hessian tangent, and that the transformation (6) has been used to make $a_{40}=a_{31}=a_{50}=a_{41}=0$.

The transformation

$$(10)\qquad x=\frac{x'}{1+Nz'},\qquad y=\frac{y'}{1+Nz'},\qquad z=\frac{z'}{1+Nz'},$$

does not change the edges $x=z=0$, $y=z=0$ of the tetrahedron of reference, and does not affect the coefficients of the forms $A_3(x, y)$, $A_4(x, y)$, $A_5(x, y)$. The new coefficient $\bar{a}_{60}$ is given by the formula

$$\bar{a}_{60}=-2Na_{30}+a_{60}.$$

Hence we may choose N in the transformation (10) in such a manner as to make $\bar{a}_{60}=0$. We shall assume that this transformation has been effected.

If we change the unit point by the transformation

$$x = px', \qquad y = qy', \qquad z = rz', \tag{11}$$

we find that, if $a_{22} \neq 0$, we may make $a_{30} = a_{12} = a_{22} = 1$, by choosing p, q, r to have the values

$$p = a_{12}/a_{22}, \qquad q = \pm (a_{12}a_{30})^{1/2}/a_{22}, \qquad r = a_{12}^3 a_{30}/a_{22}^3. \tag{12}$$

Hence *the power series expansion* (1) *may be written in the canonical form*

$$z = \alpha_3(x, y) + \alpha_4(x, y) + \alpha_5(x, y) + \cdots \tag{13}$$

wherein

$$\begin{aligned}
&\alpha_3(x, y) = x^3 + xy^2 + Ay^3, \\
&\alpha_4(x, y) = x^2y^2 + a_{13}xy^3 + a_{04}y^4, \\
&\alpha_5(x, y) = Bx^3y^2 + a_{23}x^2y^3 + a_{14}xy^4 + a_{05}y^5, \\
&\alpha_6(x, y) = Cx^5y + \cdots, \qquad \alpha_7(x, y) = Dx^7 + Ex^6y + \cdots, \\
&\alpha_8(x, y) = Fx^8 + \cdots, \qquad \alpha_9(x, y) = Gx^9 + \cdots.
\end{aligned} \tag{14}$$

The edge $y = z = 0$ is an arbitrary tangent line, not a tangent line to the curve of intersection of the surface with its tangent plane, and not a Hessian tangent. Every coefficient of the canonical expansion (13) is an absolute invariant expressing the relation between the surface S at the planar point O and the tangent line $y = z = 0$. In the expressions (14) we have written only those terms which will be of use in the discussions to follow.

If $y = z = 0$ is a Hessian tangent we may, as was remarked previously, make use of the coefficient β of the transformation (7) to make $a_{05} = 0$. We may as before choose the coefficient N of the transformation (10) to make $a_{60} = 0$. If the two Hessian tangents are not coincident, and if $a_{22} \neq 0$, by proper choice of the unit point we may make $a_{30} = a_{03} = a_{22} = 1$. Hence *the power series expansion* (1) *may be written in the canonical form*

$$\begin{aligned}
z = {} & x^3 + y^3 + x^2y^2 + A'x^4y + \cdots + B'x^5y + \cdots \\
& + C'x^7 + D'x^6y + \cdots + E'x^8 + \cdots + F'x^9 + \cdots.
\end{aligned} \tag{15}$$

The edges $y = z = 0$, $x = z = 0$ of the tetrahedron of reference are the Hessian tangent to S at O. Every coefficient of the expansion (15) is an absolute invariant of the surface.

3. Plane sections through a tangent line

Let us first consider a surface S whose equation in non-homogeneous projective coordinates (x, y, z) has been reduced to the form (13). Let us intro-

duce homogeneous projective coordinates (x_1, x_2, x_3, x_4) of a point whose non-homogeneous projective coordinates are (x, y, z) by placing

$$x = \frac{x_2}{x_1}, \qquad y = \frac{x_3}{x_1}, \qquad z = \frac{x_4}{x_1}.$$

The plane whose equation is

$$y = nz, \qquad n \neq 0, \tag{16}$$

cuts the surface in a curve Γ whose projection C onto the tangent plane from $O_3(0, 0, 0, 1)$ has the equations

$$\begin{aligned} &y = n[x^3 + xy^2 + Ay^3 + x^2y^2 + \cdots + Bx^3y^2 + \cdots + Cx^5y + \cdots], \\ &z = 0. \end{aligned} \tag{17}$$

If we solve the series in (17) for y as a power series in x we obtain

$$y = n[x^3 + b_7x^7 + b_8x^8 + b_9x^9 + \cdots] \tag{18}$$

wherein

$$b_7 = n^2 + D, \qquad b_8 = n^2 + Cn + F, \qquad b_9 = An^3 + Bn^2 + En + G. \tag{19}$$

Equations (16) *and* (18) *are the equations of an arbitrary plane section* Γ *of the surface* S *through an arbitrary tangent line not a Hessian tangent at a planar point* O *of* S.

Suppose next that the tangent line $y = z = 0$ is a Hessian tangent line, and that the equation of the surface has been reduced to the form (15). The plane (16) intersects the surface S in a curve Γ whose projection C onto the tangent plane from $O_3(0, 0, 0, 1)$ has the equations

$$y = n[x^3 + c_7x^7 + c_8x^8 + c_9x^9 + \cdots], \qquad z = 0, \tag{20}$$

wherein

$$c_7 = A'n + C', \qquad c_8 = n^2 + B'n + E', \qquad c_9 = n^3 + D'n + F'. \tag{21}$$

4. Bompiani's osculants and their loci

We may use the expansions (18) and (20) to derive the equations of the osculants used by Bompiani in discussing the various neighborhoods of an inflexion of a plane curve. We observe first that the expansions (18) and (20) are identical in form with the expansion* (2) of Bompiani's paper.

Let us first discuss the equations (18) and $z = 0$. These are the equations of the projection C of the curve Γ of intersection of the plane (16) with the surface S.

* Bompiani, p. 119.

The equations of the cusped cubic with seven-point contact with C at O are

$$y = nx^3, \qquad z = 0. \tag{22}$$

This cubic has of course an inflexion at $O(1, 0, 0, 0)$ and its cusp is at the point $O_2(0, 0, 1, 0)$. The cusp tangent of the cubic is the line $x_1 = x_4 = 0$. *The locus of the cusps for all plane sections* Γ *through* $y = z = 0$ *is therefore a line, namely, the edge* $x_1 = x_2 = 0$ *of the tetrahedron of reference. The cusp locus of all six-point cusped cubics of the curve* C *is the line* $x = z = 0$ *previously characterized as the line corresponding to the line* $y = z = 0$ *considered as a double line in the cubic involution in the pencil of tangent lines determined by* $\alpha_3(x, y) = 0$. The cusp tangent intersects the tangent $y = z = 0$ at the point $O_1(0, 1, 0, 0)$ of the tetrahedron of reference. The edges $x_1 = x_2 = 0$, $x_1 = x_4 = 0$ and the vertices O_1 and O_2 of the tetrahedron of reference giving rise to the canonical expansion (13) are therefore characterized geometrically.

The locus of the cusped cubics for all plane sections Γ is the cubic cone

$$z = x^3 \tag{23}$$

with vertex at the cusp O_2 of the seven-point cusped cubic (22).

The equations of the cubic having eight-point contact with Γ at O and passing through the cusp of the seven-point cusped cubic of Γ are

$$ny = n^2x^3 + b_7xy^2, \qquad y = nz. \tag{24}$$

The locus of the curve (24) for all sections through the given arbitrary tangent line is the algebraic surface of order three whose equation is

$$z = x^3 + x(y^2 + Dz^2). \tag{25}$$

The surfaces (23) and (25) intersect in the line $x = z = 0$, and in two plane curves whose projections from the point O_3 onto the tangent plane are the cusped cubics

$$y = \pm (-D)^{1/2}x^3, \qquad z = 0,$$

and whose projections from the point O_1 onto the face of the tetrahedron opposite O_1 are the straight lines

$$y^2 + Dz^2, \qquad x = 0.$$

The equations of the nine-point quartic curve passing through the point $O_1(0, 1, 0, 0)$, tangent to the cusp tangent $x_1 = x_4 = 0$, having a node at the point $O_2(0, 0, 1, 0)$ with the cusp locus $x_2 = x_4 = 0$ for one nodal tangent are

$$ny = n^2x^3 + b_7xy^2 + b_8x^2y^2, \qquad z = 0. \tag{26}$$

The residual nodal tangent of the quartic (26) is the line having the equations

$$b_7x_1 + b_8x_2 = 0, \qquad x_4 = 0. \tag{27}$$

The locus of the nine-point quartic curve of the section Γ for all sections Γ is the quartic surface whose equation is

$$z = x^3 + x(y^2 + Dz^2) + x^2(y^2 + Cyz + Fz^2). \tag{28}$$

The locus of the residual nodal tangent of the nine-point quartics of the sections Γ is the cubic surface

$$y^2 + Dz^2 + x(y^2 + Cyz + Fz^2) = 0. \tag{29}$$

The locus of the cusps $x_1 = x_2 = 0$ of the seven-point cusped cubics lies on the cubic surface (29). The tangent plane to the surface (29) at the point $(0, 0, \lambda, \mu)$ is the plane

$$(\lambda^2 + D\mu^2)x_1 + (\lambda^2 + C\lambda\mu + F\mu^2)x_2 = 0. \tag{30}$$

The tangent plane (30) coincides with the plane determined by the locus of the cusps and by O if and only if

$$\lambda^2 + D\mu^2 = 0.$$

There exist therefore *two points P_1, P_2 on the locus of the cusps of the seven-point cusped cubics of plane sections through an arbitrary non-Hessian tangent at a planar point O such that the tangent plane to the locus of the residual nodal tangent of the nine-point nodal quartic coincides with the plane determined by the point O and the locus of the cusps of the seven-point cusped cubics. The vertex $O_3(0, 0, 0, 1)$ is determined on the locus of the cusps as the harmonic conjugate of the cusp O_2 with respect to the points P_1, P_2.* The four vertices of the tetrahedron of reference giving rise to the expansion (13) have therefore been characterized geometrically. In the last section we shall characterize the unit point of the system of coordinates.

The seven-point cusped cubic and the nine-point quartic of the projection C of Γ intersect at O and also at the point

$$\left(-\frac{b_7}{b_8}, \; -\frac{nb_7^3}{b_8^3}, \; 0\right).$$

The coordinates of the intersection of the seven-point cusped cubic and the nine-point quartic of the curve of section Γ are

$$\left(-\frac{b_7}{b_8}, \; -\frac{nb_7^3}{b_8^3}, \; -\frac{b_7^3}{b_8^3}\right). \tag{31}$$

The locus of this point for all sections Γ of the surface through the tangent $y=z=0$ is a rational curve of order seven tangent to $y=z=0$ at O.

The line through O and the point (31) has the equations

$$b_8^2 y = nb_7^2 x, \qquad y = nz. \tag{32}$$

The locus of this line is the quintic cone

$$x(y^2 + Dz^2)^2 = z(y^2 + Cyz + Fz^2)^2. \tag{33}$$

The lines $x=z=0$, $y=z=0$ lie on the cone. The plane $x-z=0$ is tangent to the cone along the line $x=z=0$. The line $y=z=0$ is a singular line on the cone, the planes $y^2+Dz^2=0$ being tangent to the cone along this generator.

From (18) we find readily that

$$\begin{aligned} y - nx^3 &= n[b_7x^7 + b_8x^8 + \cdots], \\ y - nx^3 - \frac{b_7}{n}xy^2 &= n[b_8x^8 + b_9x^9 + \cdots], \\ y - nx^3 - \frac{b_7}{n}xy^2 - \frac{b_8}{n}x^2y^2 &= nb_9x^9 + \cdots. \end{aligned} \tag{34}$$

It follows therefore that *the seven-point cusped cubic has eight-point contact with Γ if and only if $b_7=0$, the eight-point cubic* (24) *has nine-point contact with Γ if and only if $b_8=0$, and the nine-point quartic has ten-point contact with Γ if and only if $b_9=0$. There are two plane sections through an arbitrary tangent line at a planar point of a surface such that the seven-point cusped cubic of the section has eight-point contact with the section. The planes of these sections separate the planes $y=0$, $z=0$ harmonically.* The face $y=0$ of the tetrahedron of reference giving rise to the canonical expansion (13) has therefore been given a simple geometrical characterization.

Let us now consider the osculants of Bompiani for the curves of intersection of the surface S by planes through a Hessian tangent line at the planar point. If use be made of equations (20), we find readily that the equations of the seven-point cusped cubic, the eight-point cubic, and the nine-point nodal quartic of the section Γ through a Hessian tangent are respectively

$$y = nx^3, \qquad y = nz; \tag{35}$$

$$ny = n^2x^3 + c_7xy^2, \qquad y = nz; \tag{36}$$

$$ny = n^2x^3 + c_7xy^2 + c_8x^2y^2, \qquad y = nz. \tag{37}$$

The loci of these curves are respectively

$$z = x^3, \tag{38}$$

$$z = x^3 + xz(A'y + C'z), \tag{39}$$

$$z = x^3 + xz(A'y + C'z) + x^2(y^2 + B'yz + E'z^2). \tag{40}$$

From (20) we find that

$$
\begin{aligned}
y - nx^3 &= n[c_7x^7 + c_8x^8 + \cdots], \\
y - nx^3 - \frac{c_7}{n}xy^2 &= n[c_8x^8 + c_9x^9 + \cdots], \\
y - nx^3 - \frac{c_7}{n}xy^2 - \frac{c_8}{n}x^2y^2 &= nc_9x^9 + \cdots
\end{aligned} \tag{41}
$$

It follows that *through a Hessian tangent line there is only one plane section having eight-point contact with its seven-point cusped cubic, two sections having nine-point contact with their eight-point cubics, and three sections having ten-point contact with their nine-point quartics.*

The curves (35) and (37) intersect at O and at the point

$$\left(-\frac{c_7}{c_8}, \ -\frac{c_7^3}{c_8^3}, \ -\frac{c_7^3}{c_8^3}\right).$$

The locus of the line joining O to this point is a quartic cone whose equation is

$$(y^2 + B'yz + E'z^2)^2 = xz(A'y + C'z)^2. \tag{42}$$

Other discussions similar to those made for sections of the surface through a non-Hessian tangent could be made, but we shall carry this discussion no further.

5. Loci related to the surface S at O

The points, curves, and surfaces discussed in the previous sections have, in the main, given properties of the configuration composed of the surface S and the given tangent line used in making the sections of the surface. For example, the cusp of the seven-point cusped cubic of the projection of the plane section is a point intrinsically related to the tangent line through which the plane sections were made. Likewise the locus of this point is a straight line which is related to the particular tangent line through which the sections were made.

In this section we propose to derive the necessary formulas for studying the loci of the various points, lines, and the envelopes of the surfaces derived in the previous sections as the given tangent line generates the pencil of tangent lines. Thus we obtain loci covariant to the surface S at the planar point.

To the end in view we shall make the transformation

$$x = \bar{x}, \qquad y = m\bar{x} + \bar{y}, \qquad z = \bar{z}, \tag{43}$$

on the expansion (13). We obtain an expansion of the form

$$\bar{z} = B_3(\bar{x}, \bar{y}) + B_4(\bar{x}, \bar{y}) + \cdots. \tag{44}$$

We now, by means of transformations similar to (4), (7), (10), and (11), reduce (44) to the canonical form resembling the form (13) except that the line $\bar{y}=\bar{z}=0$ is the arbitrary tangent line used as one edge of the tetrahedron of reference giving rise to the canonical expansion.

The forms $B_3(\bar{x}, \bar{y}), \cdots$ have coefficients $\bar{a}_{30}, \bar{a}_{21}, \cdots$ given by the formula

$$\bar{a}_{jk} = \frac{1}{k!} \frac{d^k \alpha_{j+k}(1, m)}{dm^k} \qquad (j + k \geqq 3). \tag{45}$$

The canonical form of the expansion (44) with $y'=z'=0$ ($\bar{y}=\bar{z}=0$) as one edge of the tetrahedron of reference is of the form

$$z' = \alpha_3'(x', y') + \alpha_4'(x', y') + \cdots . \tag{46}$$

By means of (4), (7), (10), (11), and (43) the transformation of coordinates between the tetrahedron of reference used in the canonical form (13) to that of the canonical form (46) is

$$\begin{aligned}
\rho x_1 &= x_1' + \bar{p}\overline{L}x_2' + \bar{q}\overline{M}x_3' + \bar{r}\overline{N}x_4', \\
\rho x_2 &= \bar{p}x_2' + \bar{q}\bar{\mu}x_3' + \bar{r}(\bar{\alpha} + \bar{\beta}\bar{\mu})x_4', \\
\rho x_3 &= \bar{p}m x_2' + \bar{q}(1 + m\bar{\mu})x_3' + \bar{r}[\bar{\beta} + m(\bar{\alpha} + \bar{\beta}\bar{\mu})]x_4', \\
\rho x_4 &= \bar{r}x_4',
\end{aligned} \tag{47}$$

wherein $\bar{\alpha}, \bar{\beta}, \cdots$ are defined by formulas used in reducing (1) to the form (13) but with a_{jk} replaced by $\bar{a}_{jk}$ defined by (45). These formulas are

$$\begin{aligned}
&3\bar{a}_{30}\bar{\mu} + \bar{a}_{21} = 0, \qquad 2\bar{a}_{30}\overline{L} = \bar{a}_{40}, \\
&2\bar{a}_{30}\overline{M} = \bar{a}_{31} + 4\bar{a}_{40}\bar{\mu}, \qquad 3\bar{\alpha}\bar{a}_{30}^2 = 3\overline{L}\bar{a}_{40} - \bar{a}_{50}, \\
&2\bar{a}_{30}\bar{\beta}[3\bar{a}_{30}\bar{\mu}^2 + 2\bar{a}_{21}\bar{\mu} + \bar{a}_{12}] = 3\overline{M}\bar{a}_{40} - \bar{a}_{41} - 5\bar{\mu}\bar{a}_{50}, \\
&2\bar{a}_{30}\overline{N} = \bar{a}_{60}.
\end{aligned} \tag{48}$$

Denote by O, O_1', O_2', O_3' the vertices of the tetrahedron giving rise to the canonical expansion (46) in such a manner that O_1', O_2', O_3' have geometrical characterizations similar to O_1, O_2, O_3.

By means of (47) and (48) we find readily that the parametric equations of the locus of the point O_1' as the line $y'=z'=0$ generates the pencil of tangent lines are

$$\begin{aligned}
x_1 &= m^2 + a_{13}m^3 + a_{04}m^4, \\
x_2 &= 2(1 + m^2 + Am^3), \\
x_3 &= 2m(1 + m^2 + Am^3), \\
x_4 &= 0.
\end{aligned} \tag{49}$$

The locus of O_1' of course passes through the point O_1. Moreover the tangent line to the curve at O_1 is the line $x_1=x_4=0$. Hence *the envelope of the cusp tangents of the seven-point cusped cubics of the projections C of all sections of a surface at a planar point of the third order is a rational quartic curve.* It follows also that *the plane O_1', O_2', O_3' is tangent at O_1' to the ruled surface generated by the line $O_1'O_3'$ as the tangent $y'=z'=0$ describes the pencil of tangent lines at the planar point O.*

If we eliminate m homogeneously from the parametric equations of the locus of O_1', we may write the equations of the locus of O_1' in the form

$$2(x^3 + xy^2 + Ay^3) = x^2y^2 + a_{13}xy^3 + a_{04}y^4, \qquad z = 0. \tag{50}$$

Hence *the locus of O_1' has a triple point at O, the triple point tangents being the tangents to the curve of intersection of the surface with its tangent plane at O.*

Equation (50) furnishes a rather simple interpretation of the form $\alpha_4(x, y)$. The line $x_1=x_4=0$ intersects the quartic (50) in four points. The equations of the four lines through O and these four points are

$$\alpha_4(x, y) = 0, \qquad z = 0.$$

The equations of the conic tangent to the line $y=z=0$ at O and having three-point contact at O_1 with the curve (50) are

$$x_3^2 - 2x_1x_2 = 0, \qquad x_4 = 0. \tag{51}$$

The parametric equations of the locus of the point O_2' as $y'=z'=0$ generates the pencil of tangent lines are readily found to be

$$\begin{aligned} x_1 &= 3(1 + m^2 + Am^3)(2m + 3a_{13}m^2 + 4a_{04}m^3) \\ &\quad - 4(2m + 3Am^2)(m^2 + a_{13}m^3 + a_{04}m^4), \\ x_2 &= -2(1 + m^2 + Am^3)(2m + 3Am^2), \\ x_3 &= 2(1 + m^2 + Am^3)(m^2 + 3), \\ x_4 &= 0. \end{aligned} \tag{52}$$

The locus of the cusp O_2' of the seven-point cusped cubic is therefore *a rational quintic curve. This curve has a triple point at O, the triple point tangents being the tangents to the curve of intersection of the surface with its tangent plane at O.* The tangent to the curve (52) at O_2 has the equations

$$2x_1 + 3x_2 = 0, \qquad x_4 = 0. \tag{53}$$

We have already characterized the vertices of the tetrahedron of reference giving rise to the expansion (13). We may characterize the unit point in the following manner. First the line (53) and the edges $x_1=x_4=0$, $x_2=x_4=0$ en-

able us to characterize the line $x_1-x_2=x_4=0$ by means of a cross ratio. This latter line intersects the conic (51) in two points Q_1 and Q_2. Either of these two points and the points O_1, O_2 may be used to characterize the line $x_2-x_3=x_4=0$ by means of a cross ratio. The choice of either Q_1 or Q_2 corresponds to the choice of sign in the double sign of q in (12). The point $E(1, 1, 1, 0)$ is therefore characterized. The line EO_3 intersects the cone(23) in the unit point.

MICHIGAN STATE COLLEGE,
EAST LANSING, MICH.

THE GENERALIZED THEOREM OF STOKES*

BY

STEWART S. CAIRNS

The generalized theorem of Stokes is an identity between an integral over an orientable r-manifold, M_r, and an integral over the boundary, B_{r-1}, of M_r, where M_r is on an n-space, R_n. Proofs† heretofore given have taken for R_n the space of a single coordinate system and have either assumed M_r to be analytic or imposed conditions not known to be fulfilled save by analytic manifolds. The present paper contains a proof for the case where R_n is an n-manifold of class one,‡ and (M_r, B_{r-1}) are made up, in a manner specified below, of continuously differentiable manifolds on R_n.

1. **Statement of the theorem.** An n-manifold, R_n, of class one is defined‡ by means of overlapping coordinate systems, called *allowable* systems, with regular transformations§ of class one between them. An r-cell on R_n will be called *regular*, if its closure can be parametrically defined by equations of the form

$$y_i = f_i(u_1, \cdots, u_r) \qquad (i = 1, \cdots, n), \tag{1.1}$$

where $(y) \equiv (y_1, \cdots, y_n)$ is an allowable system, the f's have continuous first partial derivatives, and the matrix $(\partial f_i/\partial u_j)$ is of rank r.

An *r-manifold*, M_r, on R_n will mean a point set whose closure, $\overline{M}_r = M_r + B_{r-1}$, is compact and connected and has the following properties: (1) Every point of M_r has an r-cell for one of its neighborhoods on $\overline{M}_r$; (2) Any point, P, on B_{r-1} has for one of its neighborhoods on $\overline{M}_r$ the closure of an r-cell with P on its boundary. We will call B_{r-1} the *boundary* of M_r. If M_r has no boundary, we refer to it as *closed*. Any point of $\overline{M}_r$ will be called *regular* if it has a regular cell (open or closed) for one of its neighborhoods on $\overline{M}_r$. The manifold M_r will be called *regular* if (1) every point of $\overline{M}_r$ is regular and

* Presented to the Society, October 29, 1932; received by the editors October 2, 1935, and, in revised form, February 26, 1936.

† For examples, see Poincaré, *Sur les résidus des intégrales doubles*, Acta Mathematica, vol. 9 (1887), pp. 321–380; Goursat, *Sur les invariants intégraux*, Journal de Mathématiques, (6), vol. 4 (1908), pp. 331–367; R. Weitzenböck, *Invariantentheorie*, 1923, Chapter XIV, §§11, 12, 14; F. D. Murnaghan, *Vector analysis and the theory of relativity*, 1922, pp. 29–33.

‡ Veblen and Whitehead, *A set of axioms for differential geometry*, Proceedings of the National Academy of Sciences, vol. 17 (1931), pp. 551–561; also *The Foundations of Differential Geometry*, Cambridge Tracts in Mathematics and Mathematical Physics, No. 29, 1932, Chapter 6.

§ Such transformations are characterized by the existence of continuous first partial derivatives and of a non-vanishing jacobian.

(2) B_{r-1} is a set of distinct closed $(r-1)$-manifolds each made up of regular i-manifolds $(i=0, \cdots, r-1)$ with the same sort of incidence relations as the i-cells of an $(r-1)$-dimensional complex.* Since any O-manifold is a point, we have here a recurrent definition of regular i-manifolds $(i=0, \cdots, n)$.

The formulation and proof of Stokes' theorem will be given for a regular manifold, M_r, on R_n. The method depends on the existence of a triangulation (σ) of $\overline{M}_r$ into regular cells. This aspect of the work is treated in two papers† by the writer. The first paper constructs the triangulation in the n-space of a single coordinate system. An extension of the construction to make it applicable on an n-manifold of class one is given in the second paper, which deals explicitly only with the case where M_r is closed.

Let $Y_{i_1 \cdots i_{r-1}}$ be an alternating tensor such that the partial derivatives $\partial(Y_{i_1 \cdots i_{r-1}})/\partial y_i$ are defined and continuous in a neighborhood of $\overline{M}_r$. Then‡

$$D_{i_1 \cdots i_r} = \left(\frac{1}{r!}\right) \delta^{\alpha_1 \cdots \alpha_r}_{i_1 \cdots i_r} \frac{\partial Y_{\alpha_2 \cdots \alpha_r}}{\partial y_{\alpha_1}} \tag{1.2}$$

is called the *Stokes tensor* of $Y_{i_1 \cdots i_{r-1}}$. Here, and throughout the paper, we apply the summation convention of tensor analysis only to Greek indices.

Taking (1.1) as the definition of a typical r-cell of (σ) and

$$y_i = g_i(v_1, \cdots, v_{r-1}) \qquad (i = 1, \cdots, n) \tag{1.3}$$

as the definition of a typical $(r-1)$-cell of (σ) on B_{r-1}, we can formulate Stokes' theorem as the identity

$$\begin{aligned} \int_{M_r} \epsilon D_{\alpha_1 \cdots \alpha_r} \frac{\partial(y_{\alpha_1} \cdots y_{\alpha_r})}{\partial(u_1 \cdots u_r)} \, du_1 \cdots du_r \\ = \pm r \int_{B_{r-1}} \epsilon' Y_{\alpha_1 \cdots \alpha_{r-1}} \frac{\partial(y_{\alpha_1} \cdots y_{\alpha_{r-1}})}{\partial(v_1 \cdots v_{r-1})} \, dv_1 \cdots dv_{r-1}, \end{aligned} \tag{1.4}$$

where (1) the integrals are to be evaluated over the separate cells of the triangulation and the results summed, and (2) the ϵ's are, on each cell, $+1$ or -1 according as the orientation of the cell by the parameters (u), or (v), agrees or disagrees with arbitrarily preassigned orientations, of M_r and B_{r-1}.

---

* O. Veblen, *Analysis Situs*, American Mathematical Society Colloquium Publications, vol. 5, 1931, pp. 76, 77.

† *On the triangulation of regular loci*, Annals of Mathematics, vol. 35 (1934), pp. 579–587; *Triangulation of the manifold of class one*, Bulletin of the American Mathematical Society, vol. 41 (1935), pp. 549–553.

‡ For the generalized Kronecker deltas, see O. Veblen, *Invariants of Quadratic Differential Forms*, Cambridge Tracts in Mathematics and Mathematical Physics, No. 24. For the tensor character of $D_{i_1 \cdots i_r}$, see the same reference or R. Weitzenböck, *Invariantentheorie*, Chapter XIV, §12.

The $\pm$ in identity (1.4) depends on the relative orientations of M_r and B_{r-1}. It is necessary to note, in addition, that the integrals in the above identity are absolute integral invariants* under transformations of parameters and under transformations between allowable coordinate systems. The factor r in identity (1.4) drops out if the summations are made, on both sides, over all combinations of the α's instead of being made as the α's run independently from one to n.

Extensions of Stokes' theorem to various loci made up of regular manifolds immediately suggest themselves.

2. **Reduction to a special case.** We note first that it is sufficient to establish identity (1.4) for a typical r-cell, σ_r, of (σ), for if the identity in this special case be applied to the sum of the r-cells of (σ), the contributions from any $(r-1)$-cell common to the boundaries of a pair of r-cells add up to zero. Since each coordinate system (y) can be interpreted as a homeomorphism between its domain and a region of euclidean n-space, we lose no generality in assuming that σ_r is a regular r-cell in the euclidean n-space of a rectangular cartesian coordinate system (y). We can, furthermore, select (y), for any σ_r, arbitrarily from all the allowable systems whose domains contain σ_r. This last possibility, together with the restrictions involved in constructing the triangulation (σ), permits us, in the course of our proof, to impose further conditions on σ_r [cf. §3(A) and §4(A) below].

3. **The generalized divergence theorem.** In the case $r=n$, the manifold M_r becomes a region of the space in which it is imbedded. Stokes' theorem, for this case, is equivalent [see §4(B) below] to the generalized divergence theorem, in which a vector field $[Y_1(y), \cdots, Y_r(y)]$ plays the role of the tensor $Y_{i_1 \cdots i_{r-1}}$ and the divergence

$$\text{div } Y = \frac{\partial Y_\alpha}{\partial y_\alpha} \tag{3.1}$$

replaces the Stokes tensor. If we regard the y's as rectangular cartesian coordinates in a euclidean space (see §2 above), we can express the divergence theorem for the region σ_r in the form

$$\int_{\sigma_r} (\text{div } Y) dV = \int_{\beta_{r-1}} Y_\alpha \gamma^\alpha d\beta, \tag{3.2}$$

where β_{r-1} is the boundary of σ_r and the γ's are the direction cosines of the outer normal to β_{r-1} at any point.

The identity (3.2) will first be established for the special case of the field

* For a proof, see R. Weitzenböck, loc. cit., §11.

$(0, \cdots, 0, Y_r)$ and a region ρ_r of the following sort. Let ρ_{r-1} be the projection of ρ_r onto $(y_1, \cdots, y_{r-1})$-space. Then the boundary, b_{r-1}, of ρ_r is made up of three parts (b^1, b^2, b^3) as follows: b^1 and b^2 are $(r-1)$-dimensional surfaces definable by equations

$$(3.3) \qquad b^j: \quad y_r = f^j(y_1, \cdots, y_{r-1}) \qquad [j = 1, 2; (y_1, \cdots, y_{r-1}, 0) \text{ on}^* \bar{\rho}_{r-1}],$$

where $f^2 > f^1$ on ρ_{r-1} and both the f's have continuous first partial derivatives on $\bar{\rho}_{r-1}$; b^3 is made up of the closed line-segments parallel to the y_r-axis which join the boundaries of b^1 and b^2. Some or all of these segments may be of length zero.

In the case of ρ_r, the divergence theorem for the field $(0, \cdots, 0, Y_r)$ reduces to

$$(3.4) \qquad \int_{\rho_r} \frac{\partial Y_r}{\partial y_r} dV = \int_{b_{r-1}} Y_r \gamma^r d\beta$$

an identity which follows, as in the proof† for three dimensions, from the equivalence of multiple and iterated integrals.

(A) We now require (see §2 above) that, for each value $j = 1, \cdots, r$, σ_r can be regarded as the sum of a finite number of distinct parts, each satisfying the above description of the region ρ_r, read with y_j in place of y_r.

This condition holds, for example, if (1) σ_r is a sufficiently close approximation to the r-simplex determined by its vertices and (2) the (y)-axes are suitably oriented. [Compare the "normal regions" of Kellogg's treatment.]†

In equation (3.4), we can now replace r by the general subscript j and (ρ_r, b_{r-1}) by (σ_r, β_{r-1}). Summing the resulting identities, we obtain the identity (3.2).

4. **The generalized theorem of Stokes.** Let $\gamma^{i_1 \cdots i_r}$ denote the direction cosines‡ of the tangent r-plane to σ_r (§2) at any point.

(A) We impose on σ_r the conditions (see §2) (1) that, for some orientation of the (y)-axes and for every set $(i_1, \cdots, i_r)$,

$$(4.1) \qquad \gamma^{i_1 \cdots i_r} \not\equiv 0 \quad \text{on} \quad \bar{\sigma}_r$$

and (2) that the projection of σ_r on the r-space of $(y_{i_1}, \cdots, y_{i_r})$ shall satisfy

---

* The symbol for a point set, modified by a bar, denotes the closure of the set.

† O. D. Kellogg, *Foundations of Potential Theory*, 1929, Chapter 4. The writer's methods are similar to those used by Kellogg in 3-space.

‡ See the writer's paper, *The direction cosines of a p-space in euclidean n-space*, American Mathematical Monthly, vol. 39 (1932), pp. 518–523. We extend the definitions there given by the convention that the direction cosines be alternating in their indices. This paper is referred to hereafter as Dir. Cos.

the restrictions [§3(A)] imposed on σ_r in the proof of the generalized divergence theorem.

We will obtain Stokes' theorem by applying the divergence theorem to each projection and summing the resulting identities.

Let $\beta^{i_1 \cdots i_{r-1}}$ be direction cosines of the tangent $(r-1)$-plane to the oriented boundary, β_{r-1}, of σ_r at any point. Then, for the euclidean case mentioned in §2, Stokes' theorem is equivalent (see Dir. Cos.) to the identity

$$\int_{\sigma_r} D_{\alpha_1 \cdots \alpha_r} \gamma^{\alpha_1 \cdots \alpha_r} d\sigma = \pm r \int_{\beta_{r-1}} Y_{\alpha_1 \cdots \alpha_{r-1}} \beta^{\alpha_1 \cdots \alpha_{r-1}} d\beta. \tag{4.2}$$

(B) *The identity* (4.2), *read for* $r=n$, *suggests the following form for the divergence theorem, where we are using the vector field of* §3, *and where the* β*'s are direction cosines of the tangent* $(r-1)$*-plane to* β_{r-1}:

$$\int_{\sigma_r} \left(\sum_{i=1}^{r} (-1)^i \frac{\partial Y_i}{\partial y_i} \right) dV = \pm \int_{\beta_{r-1}} \left(\sum_{i=1}^{r} Y_i \beta^{1 \cdots i-1, i+1 \cdots r} \right) d\beta. \tag{4.3}$$

To make the work of §3 apply to identity (4.3), we need only show that

$$\gamma^i = \pm (-1)^i \beta^{1 \cdots i-1, i+1 \cdots r} \qquad (i = 1, \cdots, r), \tag{4.4}$$

where the $\pm$ depends on the orientation of β_{r-1}. Since the numerical equality of γ^i and $\beta^{1 \cdots i-1, i+1 \cdots r}$ follows easily from geometric considerations (see Dir. Cos.), we have only to show that the signs in (4.4) are correct. Using any point P on β_{r-1} as origin, let $(u_1, \cdots, u_r)$ be a coordinate system where (1) the u_1-axis is the outer normal to β_{r-1} and (2) the $(u_2, \cdots, u_r)$-axes are on the tangent $(r-1)$-plane, L_{r-1}, to β_{r-1} and orient L_{r-1} positively. Then the agreement or disagreement in orientation between the (u)-system and the (y)-system depends on the orientation of β_{r-1}. Let

$$y_i = a_{i\alpha} u_\alpha \qquad (\alpha = 1, \cdots, r) \tag{4.5}$$

be the transformation between the y's and the u's. If A_{i1} denote the minor of a_{i1} in the determinant $|a_{ij}|$, then, since the u_1-axis is perpendicular to all the other u-axes, a value k exists such that

$$a_{i1} = (-1)^i k A_{i1}, \tag{4.6}$$

where the sign of k depends on the orientation of β_{r-1}. Since the direction cosines $(\gamma^i, \beta^{1 \cdots i-1, i+1 \cdots r})$ have the signs of (a_{i1}, A_{i1}) respectively (see Dir. Cos.), the signs in equations (4.4) are correct and our demonstration is complete.

Now let $(j_1, \cdots, j_r)$ be a fixed set of r distinct numbers from the set $(1, \cdots, n)$, and let $(m_1, \cdots, m_{n-r})$ be the complement of $(j_1, \cdots, j_r)$ with

respect to $(1, \cdots, n)$, where the m's are arranged in order of increasing magnitude. Let

$$y_{m_p} = f_{m_p}(y_{j_1}, \cdots, y_{j_r}) \qquad (p = 1, \cdots, n-r) \tag{4.7}$$

be defining equations of σ_r, where $(y_{j_1}, \cdots, y_{j_r})$ is on the projection, σ_r', of σ_r on the $\gamma^{j_1 \cdots j_r}$-plane. Applying identity (4.3) in $(y_{j_1}, \cdots, y_{j_r})$-space to the vector field*

$$Z_{j_k}(y_{j_1}, \cdots, y_{j_r}) \equiv Y_{j \cdots j_{k-1} j_{k+1} \cdots j_r}[y_{j_1}, \cdots, y_{j_r}, f_{m_1}(y_{j_1}, \cdots, y_{j_r}), \cdots, f_{m_{n-r}}(y_{j_1}, \cdots, y_{j_r})] \tag{4.8}$$

we find

$$\int_{\sigma_r'} \sum_{k=1}^{r} (-1)^k \frac{\partial Z_{j_k}}{\partial y_{j_k}} d\sigma' = \pm \int_{\beta_{r-1}'} \sum_{k=1}^{r} Z_{j_k} \bar{\beta}^{j_1 \cdots j_{k-1} j_{k+1} \cdots j_r} d\beta', \tag{4.9}$$

where β_{r-1}' is the boundary of σ_r' and hence the projection of β_{r-1}, and where the $\bar{\beta}$'s are direction cosines of the tangent $(r-1)$-plane to β_{r-1}', the orientation being determined by that of β_{r-1}. This identity is now to be interpreted in terms of the Y's, γ's, β's and integrals over σ_r and β_{r-1}.

By equation (4.8)

$$\frac{\partial Z_{j_k}}{\partial y_{j_k}} = \frac{\partial Y_{j_1 \cdots j_{k-1} j_{k+1} \cdots j_r}}{\partial y_{j_k}} + \sum_{p=1}^{n-r} \left(\frac{\partial Y_{j_1 \cdots j_{k-1} j_{k+1} \cdots j_r}}{\partial y_{m_p}} \right) \left(\frac{\partial f_{m_p}}{\partial y_{j_k}} \right) \qquad (k = 1, \cdots, r). \tag{4.10}$$

Hence

$$\sum_{k=1}^{r} (-1)^k \frac{\partial Z_{j_k}}{\partial y_{j_k}} = D_{j_1 \cdots j_r} + \sum_{k=1}^{r} \sum_{p=1}^{n-r} (-1)^k \left(\frac{\partial Y_{j_1 \cdots j_{k-1} j_{k+1} \cdots j_r}}{\partial y_{m_p}} \right) \left(\frac{\partial f_{m_p}}{\partial y_{j_k}} \right). \tag{4.11}$$

From equations (4.7) we obtain the following matrix, with rows permuted, for the positively oriented tangent r-plane to σ_r

$$(\operatorname{sgn} \gamma^{j_1 \cdots j_r}) \begin{vmatrix} 1 & 0 \cdots & 0 \\ 0 & 1 \cdots & 0 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 0 & 0 \cdots & 1 \\ \dfrac{\partial f_{m_1}}{\partial y_{j_1}} & \cdots & \dfrac{\partial f_{m_1}}{\partial y_{j_r}} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \dfrac{\partial f_{m_{n-r}}}{\partial y_{j_1}} & \cdots & \dfrac{\partial f_{m_{n-r}}}{\partial y_{j_r}} \end{vmatrix} \begin{matrix} \\ \\ \\ \\ \text{rows } (j_1, \cdots, j_r) \\ \\ \\ \\ \text{rows } (m_1, \cdots, m_{n-r}) \end{matrix}.$$

* The arguments are permuted, for convenience, on the right side of equation (4.8).

Hence (see Dir. Cos.), if $d_i^1 \cdots_{i_r}$ denote the determinant of rows $(i_1, \cdots, i_r)$ in this matrix and $\Delta = (\sum_{i_1 < \cdots < i_r} d_{i_1}^2 \cdots_{i_r})^{1/2}$, then

$$(4.12)\quad \begin{cases} \gamma^{i_1 \cdots i_r} = (\operatorname{sgn} \gamma^{i_1 \cdots i_r}) \dfrac{1}{\Delta} \\ \gamma^{i_1 \cdots i_{k-1} m_p i_{k+1} \cdots i_r} = (\operatorname{sgn} \gamma^{i_1 \cdots i_r}) \left(\dfrac{\partial f_{m_p}}{\partial y_{j_k}}\right)\left(\dfrac{1}{\Delta}\right) = \left(\dfrac{\partial f_{m_p}}{\partial y_{j_k}}\right) \gamma^{i_1 \cdots i_r}. \end{cases}$$

Therefore, substituting in equations (4.11),

$$(4.13)\quad \begin{aligned} \sum_{k=1}^{r} (-1)^k \frac{\partial Z_{i_k}}{\partial y_{j_k}} &= D_{j_1 \cdots i_r} \\ &+ \sum_{k=1}^{r} \sum_{p=1}^{n-r} (-1)^k \left(\frac{\partial Y_{j_1 \cdots j_{k-1} j_{k+1} \cdots j_r}}{\partial y_{m_p}}\right)\left(\frac{\gamma^{i_1 \cdots i_{k-1} m_p i_{k+1} \cdots i_r}}{\gamma^{i_1 \cdots i_r}}\right). \end{aligned}$$

By the definitions of direction cosines,

$$(4.14)\quad \begin{cases} d\sigma' = \pm \gamma^{i_1 \cdots i_r} d\sigma, \\ \bar{\beta}^{i_1 \cdots i_{k-1} i_{k+1} \cdots i_r} d\beta' = \beta^{i_1 \cdots i_{k-1} i_{k+1} \cdots i_r} d\beta. \end{cases}$$

Substituting from equations (4.13) and (4.14) into (4.9), we obtain

$$(4.15)\quad \begin{aligned} &\int_{\sigma_r} D_{j_1 \cdots j_r} \gamma^{i_1 \cdots i_r} d\sigma \\ &\quad + \int_{\sigma_r} \sum_{k=1}^{r} \sum_{p=1}^{n-r} (-1)^k \left(\frac{\partial Y_{j_1 \cdots j_{k-1} i_{k+1} \cdots i_r}}{\partial y_{m_p}}\right) \gamma^{i_1 \cdots i_{k-1} m_p i_{k+1} \cdots i_r} d\sigma \\ &= \pm \int_{\beta_{r-1}} \sum_{k=1}^{r} Y_{j_1 \cdots i_{k-1} i_{k+1} \cdots i_r} \beta^{i_1 \cdots i_{k-1} i_{k+1} \cdots i_r} d\beta, \end{aligned}$$

where the $\pm$ is determined by the relative orientations of σ_r and β_{r-1} and is therefore the same for all sets $(j_1, \cdots, j_r)$. When we sum all the ${}_nP_r$ such identities as (4.15), the integrand of the second integral on the left becomes

$$(4.16)\quad \sum_{j_1 \cdots j_r} \sum_{k=1}^{r} \sum_{p=1}^{n-r} (-1)^k \left(\frac{\partial Y_{j_1 \cdots j_{k-1} i_{k+1} \cdots j_r}}{\partial y_{m_p}}\right) \gamma^{i_1 \cdots i_{k-1} m_p i_{k+1} \cdots i_r}.$$

Now let $(s_1, \cdots, s_r)$ be a fixed subset of $(1, \cdots, n)$. In the summation (4.16), the superscripts of γ become $(s_1, \cdots, s_r)$ whenever

$$(4.17)\quad \begin{cases} \text{(a)} & j_i = s_i \qquad (i \neq k), \\ \text{(b)} & j_k \notin (s_1, \cdots, s_r), \\ \text{(c)} & m_p = s_k. \end{cases}$$

For each value of k, there are $(n-r)$ sets $(j_1, \cdots, j_r)$ satisfying (4.17a) and (4.17b). Corresponding to each such set of j's and value of k, there is just one value of p satisfying (4.17c). Hence the triple summation (4.16) reduces to

$$(4.18)\qquad (n-r)\sum_{s_1\cdots s_r}\left[\sum_{k=1}^{r}(-1)^k\frac{\partial Y_{s_1\cdots s_{k-1}s_{k+1}\cdots s_r}}{\partial y_{s_k}}\right]\gamma^{s_1\cdots s_r} = (n-r)D_{\alpha_1\cdots\alpha_r}\gamma^{\alpha_1\cdots\alpha_r},$$

and the left side of identity (4.15) becomes

$$(4.19)\qquad (n-r+1)\int_{\sigma_r} D_{\alpha_1\cdots\alpha_r}\gamma^{\alpha_1\cdots\alpha_r}d\sigma.$$

The term $Y_{t_1\cdots t_{r-1}}\beta^{t_1\cdots t_{r-1}}$, where the t's are a fixed subset of $(1, \cdots, n)$, appears in the identity (4.15) if and only if $(t_1, \cdots, t_{r-1})$ can be obtained from $(j_1, \cdots, j_r)$ by deleting one of the j's. There are r possible positions in the set $(j_1, \cdots, j_r)$, for the j which is to be deleted, and $(n-r+1)$ possible values. Hence $Y_{t_1\cdots t_{r-1}}\beta^{t_1\cdots t_{r-1}}$ appears in $(n-r+1)r$ identities such as (4.15). When we sum all the identities such as (4.15), we therefore find

$$(4.20)\qquad (n-r+1)\int_{\sigma_r} D_{\alpha_1\cdots\alpha_r}\gamma^{\alpha_1\cdots\alpha_r}d\sigma = \pm\,(n-r+1)r\int_{\beta_{r-1}} Y_{\alpha_1\cdots\alpha_{r-1}}\beta^{\alpha_1\cdots\alpha_{r-1}}d\beta$$

which is equivalent to identity (4.2). As remarked in §2, the establishment of this identity completes our proof of the generalized theorem of Stokes.

LEHIGH UNIVERSITY,
BETHLEHEM, PA.

SUMMATION OF MULTIPLE FOURIER SERIES BY SPHERICAL MEANS*

BY

SALOMON BOCHNER

PART I. AN OUTLINE OF THE RESULTS

Let $x=(x_1, \cdots, x_k)$ be a point in the k-dimensional Euclidean space and $f(x)=f(x_1, \cdots, x_k)$ a function of the Lebesgue class L having the period 2π in each variable; let

$$f(x) \sim \sum a_{n_1 \cdots n_k} e^{i(n_1 x_1 + \cdots + n_k x_k)} \tag{1}$$

be its Fourier series,

$$a_{n_1 \cdots n_k} = \frac{1}{(2\pi)^k} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} f(x) e^{-i(n_1 x_1 + \cdots + n_k x_k)} dx. \tag{2}$$

We shall consider criteria for (convergence and) summability not of the rectangular partial sums

$$\sum_{-N_1}^{N_1} \cdots \sum_{-N_k}^{N_k} a_{n_1 \cdots n_k} e^{i(n_1 x_1 + \cdots + n_k x_k)} \tag{3}$$

but of the spherical partial sums

$$S_R(x) = \sum_{\nu \leqq R} a_{n_1 \cdots n_k} e^{i(n_1 x_1 + \cdots + n_k x_k)}; \qquad \nu^2 = n_1^2 + \cdots + n_k^2. \tag{4}$$

This means that we shall consider the series (1) as the simple series

$$\sum_{j=0}^{\infty} \sum_{\nu = R_j} a_{n_1 \cdots n_k} e^{i(n_1 x_1 + \cdots + n_k x_k)}, \tag{5}$$

where R_j^2 is the sequence of all integers that can be represented as a sum of k squares of integers. (For $k \geqq 4$ the sequence consists of all non-negative integers.) In summing series (5) we shall restrict ourselves to Toeplitz matrices $\{a_{pq}\}$ in which $a_{pq}=\phi(p/q)$. Thus we shall consider, and take the limit of, as $R \to \infty$, the "partial sums"

$$S_R^{\phi}(x) \sim \sum \phi\left(\frac{\nu}{R}\right) a_{n_1 \cdots n_k} e^{i(n_1 x_1 + \cdots + n_k x_k)} \tag{6}$$

---

* Presented to the Society, April 20, 1935; received by the editors October 11, 1935.

A brief summary of the results was published in the Proceedings of the National Academy of Sciences, vol. 21 (1935), pp. 353–355.

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

(the sum to be extended over all combinations of integers $n_1, \cdots, n_k$), where $\phi(t)$ is a fixed function in $0 \leqq t < \infty$ for which $\phi(0) = 1$. We note that we shall allow the index R to range over all values $0 \leqq R < \infty$ and not only over the sequence $\{R_j\}$.

Interesting cases arise when $\phi(t)$ is one of the following functions:

(7) e^{-t} Poisson-Abel summation,

(8) e^{-t^2}, Gauss-Sommerfeld summation,

(9) $$K_\delta(t) = \begin{cases} (1-t^2)^\delta & \text{for } 0 \leqq t < 1, \\ 0 & \text{for } 1 \leqq t, \end{cases}$$

the latter being the important case of Riesz summation of order δ.*

The spherical partial sums (6) have a great advantage over the rectangular partial sums (3). This will be evident from the formula

$$S_R^\phi(x) = R \int_0^\infty f_x(t) H_\phi(tR)\, dt \tag{10}$$

which we are going to explain presently (its validity will be discussed later on). We shall use throughout the abbreviations

$$l = \frac{k-2}{2}, \tag{11}$$

$$\alpha = \frac{k-1}{2}, \tag{12}$$

$$c = tR, \tag{13}$$

and by $J_l(u)$ we shall denote the Bessel function of order l. The kernel $H_\phi(c)$ depends only on $\phi(t)$ and not on $f(x)$:

$$H_\phi(c) = \frac{1}{c} \int_0^\infty \phi\left(\frac{u}{c}\right) u^{l+1} J_l(u)\, du \tag{14}$$

$$= c^{l+1} \int_0^\infty \phi(u) u^{l+1} J_l(uc)\, du. \tag{15}$$

Thus

$$H_\phi(c) = \frac{2^{l+1}\Gamma(l+\frac{3}{2})}{\pi^{1/2}} \cdot \frac{c^{2l+1}}{(1+c^2)^{l+3/2}} \quad \text{if } \phi(t) = e^{-t}, \tag{16}$$

$$H_\phi(c) = \frac{c^{2l+1}}{2^{l+1}} \cdot e^{-c^2/4} \quad \text{if } \phi(t) = e^{-t^2}, \tag{17}$$

* It is more convenient to consider the Riesz factor $(1-t^2)^\delta$ than $(1-t)^\delta$; see formula (18).

$$(18)\quad H_\phi(c) \equiv H_\delta(c) = 2^\delta\Gamma(\delta+1)\cdot\frac{J_{\delta+l+1}(c)}{c^{\delta-l}} \qquad \text{if } \phi(t) = K_\delta(t),^*$$

and, in particular,

$$(19)\quad H_\alpha(c) = 2^\alpha\Gamma(\alpha+1)\,\frac{J_{k-1/2}(c)}{c^{1/2}} \qquad \text{if } \phi(t) = K_\alpha(t).$$

The factor $f_x(t)$ in the integral (10) is determined by the given function $f(x)=f(x_1, \cdots, x_k)$. Unlike $f(x)$ it is a function of one variable only; its argument t ranges over the interval $0 \leqq t < \infty$. Its value is, apart from a factor depending only on k, the $(k-1)$-dimensional spherical mean of the function $f(x)$ over a sphere whose radius is t and whose center is at the point x. If σ denotes the unit sphere $\xi_1^2 + \cdots + \xi_k^2 = 1$ and $d\sigma_\xi$ its $(k-1)$-dimensional volume element, then

$$(20)\quad f_x(t) = \frac{1}{(2\pi)^{k/2}}\int_\sigma f(x_1 + t\xi_1, \cdots, x_k + t\xi_k)d\sigma_\xi.$$

For $k=1$ the integral over the unit sphere is to be replaced by $f(x+t)+f(x-t)$. We shall require the explicit value of $f_x(t)$ only for

$$(21)\quad f(x) = e^{i(n_1x_1+\cdots+n_kx_k)}.$$

In this case we have

$$(22)\quad f_x(t) = f_0(t)f(x).$$

Moreover

$$(23)\quad f_0(t) = \frac{J_l(\nu t)}{(\nu t)^l}.$$

In fact, it is easily seen that

$$f_0(t) = \frac{\omega_{k-1}}{(2\pi)^{k/2}}\int_0^\pi e^{i\nu t\cos\theta}(\sin\theta)^{k-2}d\theta,$$

where ω_k is the $(k-1)$-dimensional volume of the unit sphere $\xi_1^2 + \cdots + \xi_k^2 = 1$. Now apply formula (6) in Watson, loc. cit., p. 48.

As soon as formula (10) holds, what decides the convergence of the partial sums (6) at a given point x are only the spherical averages of the functions $f(x)$ around the one point x and nothing else. It must be pointed out that formula (10), as it stands, involves the function $f_x(t)$ in its entire infinite in-

* See G. N. Watson, *Bessel Functions*, p. 386, formula (6); p. 394, formula (4); p. 373, formula (1).

terval $0 \leqq t < \infty$ and not only in some such finite t-interval that might cover a fundamental domain of periodicity of $f(x)$. However, if $\phi(t)$ has a sufficient number of derivatives which are sufficiently small at infinity or, in case $\phi(t) = K_\delta(t)$, if $\delta > \alpha$, then the *limit behaviour* of $S_R^\phi(x)$, that is, its convergence or non-convergence as $R \to \infty$, depends only on the values of $f_x(t)$ in any interval $0 \leqq t < t_0$, no matter how small; in particular the partial sum $S_R^\phi(x)$ converges to 0 if the function $f(x)$ vanishes in some neighborhood of the point x (Theorem VI).

This localization property is the more remarkable since it is highly characteristic of our spherical method of summation. In the case of "rectangular" or even "square" partial sums, that is, in case we consider

$$(24) \qquad \sigma_R^\phi(x) \sim \sum \phi\left(\frac{|n_1|}{R}\right) \cdots \phi\left(\frac{|n_k|}{R}\right) a_{n_1 \cdots n_k} e^{i(n_1 x_1 + \cdots + n_k x_k)},$$

it fails to hold for such a smooth summation function as $\phi(t) = e^{-t}$ is (Theorem VIII).

As far as our Theorems I and V are concerned, a function $f(x)$ has a "limit" at a point x as soon as $f_x(t)$ is continuous for $t = 0$. In the case of rectangular summation the corresponding condition requires continuity in "octants," that is, the existence of the 2^k limits $f(x_1 \pm 0, \cdots, x_k \pm 0)$. This condition is rather more restrictive. The nearest approach to our limit "in spherical average" is the one occurring in C. N. Moore's criterion for summation of "restricted" rectangular partial sums. The condition also requires continuity in "octants," but the octants may be bounded by very general curves instead of by straight lines.*

In order to safeguard the localization property under Riesz summation it is not only sufficient but also necessary to require that $\delta > \alpha$. That the localization property no longer holds for $\delta < \alpha$ suggests itself readily and can be proved rather easily. The situation is less obvious in the case of the critical exponent $\delta = \alpha$ itself. For $k = 1$, in which case $\alpha = 0$, it is a fundamental result of Riemann that the localization property holds for $\delta = \alpha$. But, as we shall see, the contrary is true for every $k \geqq 2$. If $k \geqq 2$ then there exists a function $f(x)$ of Lebesgue class L vanishing in the neighborhood of the point $0 = (0, \cdots, 0)$ for which the partial sums $S_R^\alpha(0)$ are not bounded as $R \to \infty$ (Theorem VII).

However, the localization property does hold also for the critical ex-

* C. N. Moore, *On the summability of the double Fourier series of discontinuous functions*, Mathematische Annalen, vol. 74 (1913), pp. 555–572; esp. p. 567, Theorem III.

ponent $\delta=\alpha$ if instead of Fourier series we consider Fourier integrals. Let $f(x)$ be a (non-periodic) function belonging to class L over the whole space. Replacing (1), (2), (6) by

$$(25)\quad f(x)\sim\int\cdots\int_{-\infty}^{\infty}a(n_1,\cdots,n_k)e^{i(n_1x_1+\cdots+n_kx_k)}dn_1\cdots dn_k$$

$$(26)\quad a(n_1,\cdots,n_k)=\frac{1}{(2\pi)^k}\int\cdots\int_{-\infty}^{\infty}f(x)e^{-i(n_1x_1+\cdots+n_kx_k)}dx_1\cdots dx_k$$

$$(27)\quad S_R^{\phi}(x)\sim\int\cdots\int_{-\infty}^{\infty}\phi\left(\frac{\nu}{R}\right)a(n_1,\cdots,n_k)e^{i(n_1x_1+\cdots+n_kx_k)}dn_1\cdots dn_k$$

the relations (10)–(20) remain valid without any change whatsoever. Many results are analogous for the two groups of expansions, and in fact the Fourier integrals are easier than the Fourier series. A study of spherical summation in case of Fourier integrals was made by the author some time ago,* and in the present paper we shall give a more detailed account.

The spherical average $f_x(t)$ and the critical exponent (12) for Riesz summation of multiple Fourier series are not new. They have been known with reference to expansions in ultra-spherical functions.

The elementary exponentials

$$(28)\qquad u(x)=e^{i(n_1x_1+\cdots+n_kx_k)}$$

(all $n_1,\cdots,n_k$ integers) are a complete set of regular solutions of the characteristic value problem

$$(29)\qquad \Delta u(x)=-\lambda u(x),$$

if this equation is being considered on the (closed) torus

$$0\leqq x_1<2\pi,\cdots,0\leqq x_k<2\pi,$$

and $\Delta u(x)$ is the Laplace operator with respect to the Euclidean metric on the torus, namely,

$$\frac{\partial^2u}{\partial x_1^2}+\cdots+\frac{\partial^2u}{\partial x_k^2}.$$

Since $\lambda=n_1^2+\cdots+n_k^2$, our way of writing series (1) in the form (5) satisfies the very natural principle of ordering the terms in series (1) according to the magnitude of the characteristic values λ. Now the characteristic

* See S. Bochner, *Ein Konvergenzsatz für mehrvariablige Fouriersche Integrale*, Mathematische Zeitschrift, vol. 34 (1931), pp. 440–447; analogous ideas are found also in earlier work of A. C. Berry, *The Fourier transform theorem*, Journal of Mathematics and Physics, Massachusetts Institute of Technology, vol. 8 (1929), pp. 106–118, and N. Wiener, *Laplacians and continuous linear functionals*, Acta Szeged, vol. 3 (1927), pp. 7–16.

value problem (29) may be treated on any closed k-dimensional space with a regular Riemann metric, if $\Delta u(x)$ is understood to be the corresponding Laplace-Beltrami operator. Hilbert has shown (explicitly for $k=2$)* that in all such cases the solutions of (29) form a complete set of functions on the space. In case the given space has constant curvature the solutions of (29) are essentially the ultra-spherical functions. Expansions in ultra-spherical functions have been studied by many authors;† in the theory of Riesz summation, the exponent (12) again plays a decisive role. But the situation is slightly different. For expansions in ultra-spherical functions the localization property holds strictly only if $\delta \geqq 2\alpha$, whereas if $2\alpha > \delta > \alpha$ the antipode of the given point x has also some influence.

The similarity of results relating to Fourier integrals and Fourier series suggests the investigation of larger classes of expansions. It would not be very profitable to include generalized Fourier integrals, but it will be rather illuminating to consider Fourier series not only of periodic functions but of the larger class of almost periodic functions. On this larger basis it will be easier to understand why the Riesz summation of order (12) no longer obeys the localization property in case $k \geqq 2$, and why all the trouble with the "cross-neighborhood" of the given point x which figures so prominently in the theory of summation by rectangular means is entirely eliminated in the case of spherical summation.

The almost periodic functions we shall deal with will be the functions belonging to a Stepanoff class S_p, $p \geqq 1$.‡ For our purposes we may characterize these functions as follows. The function $f(x)$ belongs to S_p if it belongs to the Lebesgue class L_p on every finite interval, and if there exists a sequence of exponential polynomials $f^n(x)$, that is, of functions each of which is a *finite* sum of the form

$$\sum a^{(n)}_{n_1 \cdots n_k} e^{i(n_1 x_1 + \cdots + n_k x_k)}$$

(the numbers $n_1, \cdots, n_k$ being *arbitrary* real numbers), such that, for every k-dimensional unit-sphere E of radius 1, and uniformly in *all* such spheres,

$$\lim_{n \to \infty} \int \cdots \int_E | f(x) - f^n(x) |^p dx_1 \cdots dx_k = 0. \tag{30}$$

* D. Hilbert, *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen*, 1912, Chapter XVIII. Grouping of terms by spherical shells suggests itself naturally in multiple Fourier series and other expansions arising in mathematical physics, for instance, in the discussion of vibrations of rectangular membranes. See Rayleigh, *Theory of Sound*, vol. I, 1894, Chapter IX.

† See E. Kogbetliantz, *Sommation des séries et intégrales divergentes par les moyennes arithmétiques et typiques*, Mémorial des Sciences Mathématiques, Fasc. 51, 1931, pp. 70–73.

‡ See A. S. Besicovitch, *Almost Periodic Functions*, Cambridge, 1932.

An almost periodic function determines, and is determined by a Fourier series of the form (1), the numbers $n_1, \cdots, n_k$ being arbitrary real numbers.

We shall see that the formulas and results concerning spherical summation extend very naturally to almost periodic functions in general. This explains the absence of conditions involving "cross-neighborhoods." Such conditions imply discriminations between the various directions in the underlying space. But no such discriminations are likely to play a role with almost periodic functions, since the class of these functions is invariant with respect to affine transformations of the x-space.

PART II. FOURIER INTEGRALS

In the present section $f(x)$ is a measurable function for which

$$\int \cdots \int_{-\infty}^{\infty} | f(x) | \, dx_1 \cdots dx_k < \infty \tag{31}$$

and $S_R^{\phi}(x)$ is defined by (25)–(27).

The function $f_x(t)$, which was defined by (20), is measurable, and

$$\int_0^{\infty} | f_x(t) | \, t^{k-1} dt < \infty . \tag{32}$$

LEMMA 1. *The formula*

$$S_R^{\phi}(x) = R \int_0^{\infty} f_x(t) H_{\phi}(tR) dt \tag{10}$$

holds if

$$\int_0^{\infty} | \phi(t) | \, t^{k-1} dt < \infty . \tag{33}$$

In the proof we may assume $x = (0, \cdots, 0)$. Substituting (26) in the relation (27) and changing the order of integrations (which is justified by (33)) we obtain

$$S_R^{\phi}(0) = \int \cdots \int_{-\infty}^{\infty} f(x) F(x) dx_1 \cdots dx_k, \tag{34}$$

where

$$F(x) = \frac{1}{(2\pi)^k} \int \cdots \int_{-\infty}^{\infty} \phi\left(\frac{\nu}{R}\right) e^{-i(n_1 x_1 + \cdots + n_k x_k)} dx_1 \cdots dx_k .$$

By an established formula*

* S. Bochner, *Vorlesungen über Fouriersche Integrale*, 1932, p. 186, Satz 56.

$$(2\pi)^{k/2}F(x_1, \cdots, x_k) = t^{-l}\int_0^\infty \phi\left(\frac{u}{R}\right)u^{l+1}J_l(ut)du$$
$$= Rt^{-k+1}H_\phi(tR),$$

where $t^2 = x_1^2 + \cdots + x_k^2$. Substituting this in (34) and making the change of coordinates $x_1 = t\xi_1, \cdots, x_k = t\xi_k$ we finally obtain

$$S_R^\phi(0) = R\int_0^\infty f_0(t)H_\phi(tR)dt. \tag{35}$$

LEMMA 2. *Let $\phi(t)$ be either the Riesz kernel $K_\delta(t)$ with $\delta > \alpha$, or a (complex) function in $0 \leqq t < \infty$ having the following properties*:

(i) *The inequality*

$$\int_0^\infty |\phi(t)|\, t^{k-1}dt < \infty \qquad (k - 1 = 2l + 1) \tag{36}$$

holds.

(ii) *If g is the integer defined by*

$$\tfrac{1}{2} > l - g \geqq -\tfrac{1}{2}, \tag{37}$$

*then $\phi(t)$ has $(g+2)$ derivatives in $0 \leqq t < \infty$, each bounded in some neighborhood of $t = 0$, such that**

$$\overline{\lim_{u\to\infty}}\, |\phi^{(\gamma)}(u)u^\lambda| < \infty, \qquad \int_0^\infty |\phi^{(\gamma)}(u)u^\lambda|\, du < \infty \tag{38}$$

for

$$\gamma = 0, 1, \cdots, g + 2, \qquad 0 \leqq \lambda \leqq l + \tfrac{1}{2}.$$

Under these circumstances, $H_\phi(c)$ has the following properties:

$$|H_\phi(c)| \leqq A_1 c^{k-1} \tag{39}$$

$$|H_\phi(c)| \leqq \frac{A_2}{c^{1+\kappa}} \tag{40}$$

$$R\int_0^\infty H_\phi(tR)dt = \phi(0)2^l\Gamma(l + 1), \tag{41}$$

where A_1, A_2, κ are positive constants independent of c.

For $\phi(t) = K_\delta(t)$ we have (18); (39) and (40) follow immediately, since $J_\nu(c) = O(c^\nu)$ as $c \to 0$, and $= O(c^{-1/2})$ as $c \to \infty$. As to (41), see Watson, loc. cit., p. 391, formula (1).

* The conditions (38) are redundant for some values of γ and λ.

Now let $\phi(t)$ satisfy hypothesis (i). Since $|J_l(uc)| \leqq A(uc)^l$, we have

$$\left|\int_0^\infty \phi(u)u^{l+1}J_l(uc)du\right| \leqq Ac^l\int_0^\infty |\phi(u)|\,u^{2l+1}du,$$

and this proves (39). As to (40), what we have to prove is

$$\int_0^\infty \phi\left(\frac{u}{c}\right)u^{l+1}J_l(u)du = O(c^{-\kappa}), \qquad c\to\infty. \tag{42}$$

We shall prove more, namely,

LEMMA 3. *If l is any real number $\geqq -\frac{1}{2}$, and if a function $\phi(t)$ satisfies hypothesis* (ii) *of Lemma* 2, *then* (42) *holds.*

We shall first treat the case $\frac{1}{2}>l\geqq -\frac{1}{2}$, that is, $g=0$. We observe that $u^{l+1}J_l(u)$ is the derivative of $u^{l+1}J_{l+1}(u)$ and that the function

$$P(u) = \int_0^u u^{l+1}J_{l+1}(u)du$$

has the properties

$$P(u) = O(u^{2l+3}) = O(u^{l+1/2}) \qquad u\to 0,$$
$$P(u) = O(u^{l+1/2}) \qquad u\to\infty.$$

Making use of the various relations constituting (38) we can readily justify the following integrations by parts:

$$\int_0^\infty \phi\left(\frac{u}{c}\right)u^{l+1}J_l(u)du = -\frac{1}{c}\int_0^\infty \phi'\left(\frac{u}{c}\right)u^{l+1}J_{l+1}(u)du$$
$$= \frac{1}{c^2}\int_0^\infty \phi''\left(\frac{u}{c}\right)P(u)du.$$

Majorizing $P(u)$ by $u^{l+1/2}$ in the whole interval $(0, \infty)$, we obtain (42) with $\kappa=\frac{1}{2}-l$. This disposes of the case $g=0$. The case of an arbitrary integer $g>0$ can now be dealt with by a process of induction, the induction working from the pair of values (g, l) to the pair of values $(g+1, l+1)$. In fact,

$$\int_0^\infty \phi\left(\frac{u}{c}\right)u^{l+2}J_{l+1}(u)du = \int_0^\infty \phi\left(\frac{u}{c}\right)u^{2l+2}\frac{J_{l+1}(u)}{u^l}du$$
$$= \int_0^\infty \frac{d}{du}\left(\phi\left(\frac{u}{c}\right)u^{2l+2}\right)\frac{J_l(u)}{u^l}du = \int_0^\infty \psi\left(\frac{u}{c}\right)u^{l+1}J_l(u)du,$$

where

$$\psi(u) = (2l+2)\phi(u) + u\phi'(u);$$

and the condition (38) is so arranged that if $\phi(u)$ satisfies it for (g_0+1, l_0+1), $\psi(u)$ satisfies it for (g_0, l_0). This completes the proof of Lemma 3, and therefore, the proof of (40).

Relation (41) will follow from the Fourier-Hankel inversion formula* which if applied to the function $\phi(u)u^{l+1/2}$ reads

$$\int_0^\infty \frac{J_l(cx)}{(cx)^l}\,dc \int_0^\infty \phi(u)(uc)^{l+1}J_l(uc)du = \phi(x), \tag{43}$$

or

$$\int_0^\infty \frac{J_l(cx)}{(cx)^l} H_\phi(c)dc = \phi(x). \tag{44}$$

Since $J_l(cx)\cdot(cx)^{-l}$ is boundedly convergent if $x\to 0$, and $H_\phi(c)$ is absolutely integrable in $(0, \infty)$ on account of (39) and (40), we may deduce from (44) the relation

$$\int_0^\infty H_\phi(c)dc = \phi(0)\cdot 2^l\Gamma(l+1), \tag{45}$$

which is the same as (41).

THEOREM I. *Let $\phi(t)$ be a function as in Lemma 2 for which $\phi(0)=1$. If for a given function $f(x)$ and a fixed point x there exists the limit*

$$\lim_{t\to 0} f_x(t) = s,$$

or, more generally, if

$$\lim_{t\to 0} t^{-k}\int_0^t t^{k-1}\left|f_x(t)-s\right|dt = 0, \tag{46}$$

then

$$\lim_{R\to\infty} S_R^\phi(x) = 2^l\Gamma(l+1)\cdot s. \tag{47}$$

From (10) and (41) follows

$$S_R^\phi(x) - 2^l\Gamma(l+1)\cdot s = R\int_0^\infty g(t)H_\phi(tR)dt, \tag{48}$$

where

$$g(t) = f_x(t) - s. \tag{49}$$

* See Watson, loc. cit., p. 456, §144.

In (48) we split the interval of integration in $(0, R^{-1})$ and (R^{-1}, ∞). Using (39) for $tR \leqq 1$ and (40) for $tR \geqq 1$, we see that (48) may be majorized by the sum of

$$R^k F(R^{-1}) \tag{50}$$

and

$$R^{-\kappa} \int_{R^{-1}}^{\infty} t^{-\kappa-k} dF(t), \tag{51}$$

where

$$F(t) = \int_0^t | g(t) | t^{k-1} dt.$$

By (46) and (32),

$$\begin{aligned} F(t) &= o(t^k), \quad \text{as } t \to 0, \\ &= O(t^k), \quad \text{as } t \to \infty ; \end{aligned} \tag{52}$$

from this it follows easily that (50) and (51) both converge to 0 as $R \to \infty$.

Theorem I covers the Riesz kernel $\phi(t) = K_\delta(t)$ only for $\delta > \alpha$. Now let $\delta = \alpha - \beta$, $0 \leqq \beta < 1$. Since, up to a multiplicative constant, $H_\phi(t)$ has the value

$$c^{l-\delta} J_{\delta+l+1}(c), \tag{53}$$

we again have the relations (39) and (41), the integral in the latter relation being conditionally convergent at ∞; whereas instead of (40) we shall use, for $tR \geqq 1$, the relation

$$H_\delta(tR) = A_{k\delta} R^{\beta-1} \frac{\cos (tR - (k + \beta)\pi/2)}{t^{1-\beta}} + O((tR)^{\beta-2}) \tag{54}$$

in which $A_{k\delta}$ is independent of t and R. By Lemma 1 in connection with (41) we again have relation (48) and we again put

$$R \int_0^{\infty} g(t) H_\phi(tR) dt = R \int_0^{R^{-1}} + R \int_{R^{-1}}^{\infty} = J_1(R) + J_2(R).$$

On account of (39) the mere condition (46) is again sufficient to secure the relation $\lim_{R \to \infty} J_1(R) = 0$. But the previous estimation of $J_2(R)$ no longer holds true. If we substitute the value (54) for $H_\delta(tR)$ in $J_2(R)$ then the previous estimation applies only to the contribution arising from the error term. Thus we obtain the following

THEOREM II. *Let* $\delta=\alpha-\beta$, $0\leqq\beta<1$. *If* $f_x(t)$ *satisfies condition* (46), *then the limit relation*

$$\lim_{R\to\infty} S_R^{\delta}(x) = 2^l\Gamma(l+1)\cdot s \tag{55}$$

holds if, and only if,

$$\lim_{R\to\infty} R^{\beta}\int_{R^{-1}}^{\infty}(f_x(t)-s)\frac{\cos(tR-(k+\beta)\pi/2)}{t^{1-\beta}}\,dt = 0. \tag{56}$$

For any $\epsilon>0$ and $0\leqq\beta<1$,

$$R^{\beta}\int_{\epsilon}^{\infty}\frac{\cos(tR-(k+\beta)\pi/2)}{t^{1-\beta}}\,dt = \int_{\epsilon R}^{\infty}\frac{\cos(u-(k+\beta)\pi/2)}{u^{1-\beta}}\,du \to 0 \tag{57}$$

as $R\to\infty$.

On the other hand, for $\beta>0$, the assumption (32) does *not* imply

$$R^{\beta}\int_{\epsilon}^{\infty} f_x(t)\frac{\cos(tR-(k+\beta)\pi/2)}{t^{1-\beta}}\,dt \to 0 \qquad \text{as } R\to\infty. \tag{58}$$

In fact, for $\beta>0$, (58) need not hold even for a function $f(x)$ that vanishes outside a finite interval. However, for $\beta=0$, it is immediately seen that (32) implies (58) for $k>1$. Combining this with the familiar result in case $k=1$, we obtain the following

THEOREM III. *If* $f_x(t)$ *satisfies condition* (46) *then the limit relation*

$$\lim_{R\to\infty} S_R^{\alpha}(x) = 2^l\Gamma(l+1)\cdot s \tag{59}$$

holds if, and only if, for some (or every) $\epsilon>0$,

$$\lim_{R\to\infty}\int_{R^{-1}}^{\epsilon}(f_x(t)-s)\frac{\cos(tR-k\pi/2)}{t}\,dt = 0. \tag{60}$$

Thus for the validity of (59) *it is sufficient that* $g(t)=f_x(t)-s$ *satisfies the conditions of Lebesgue*

$$\int_0^{\eta}|g(t)|\,t^{k-1}dt = o(\eta^k), \qquad \int_{\eta}^{\epsilon}\frac{|g(t)-g(t+\eta)|}{t}\,dt \to 0 \quad \text{as } \eta\to 0. \tag{61}$$

For $\delta\geqq\alpha$, *the relation* (55) *depends only on the behaviour of* $f(x)$ *in a neighborhood of the given point* x; *this is no longer the case for* $\delta<\alpha$.

For $\delta<\alpha$, criteria for the validity of (55) will of necessity involve the behaviour of $f_x(t)$ in its entire interval $0\leqq t<\infty$. For $0<\beta<1$, if (46) holds, it

is sufficient to find conditions for the validity of (56). By classical methods it is easy to find that, using (46), (56) will hold if

$$R^{\beta}\int_{R^{-1}}^{\infty}\frac{|g(t)-g(t+\pi/R)|}{t^{1-\beta}}dt\to 0 \qquad \text{as } R\to\infty .$$

This in its turn, will be satisfied if, for instance, $g(t)$ is the Lebesgue integral of a function $g'(t)$ in every finite interval $0\leqq t\leqq t_0$ such that

$$\int_1^{\infty}\frac{|g'(t)|}{t^{1-\beta}}dt<\infty .$$

$g(t)$ thus being absolutely continuous in $0\leqq t\leqq t_0$, this condition implies the existence of the limit $g(+0)$, which is much more special than (46). Altogether we obtain the following criterion. *For $0<\beta<1$, (55) will hold, and s will be $f_x(+0)$, if $f_x(t)$ is absolutely continuous in every finite interval $0\leqq t\leqq t_0$, and*

$$\int_1^{\infty}\frac{|f_x'(t)|}{t^{1-\beta}}dt<\infty .$$

For $\delta<\alpha-1$, it is easy to find criteria for the validity of (55) involving the existence of derivatives of higher than the first order for the function $f_x(t)$. But there would not be much to be gained from a detailed analysis. The interesting case $\delta=0$ has been treated elsewhere.*

PART III. FOURIER SERIES

LEMMA 4. *Corresponding to any dimension number k, any $p\geqq 1$, and any $\epsilon>0$, there exists a number $A=A(k, p, \epsilon)$ with the following property: if $f(x)=f(x_1, \cdots, x_k)$ belongs to class L_p over every finite interval, and if there exists a number B such that*

$$\int_E |f(x)|^p dx_1\cdots dx_k\leqq B \tag{62}$$

for every k-dimensional sphere E of radius 1, then, for every point x,

$$\int_t^{t+1}|f_x(t)|^p dt\leqq AB, \qquad t\geqq\epsilon, \tag{63}$$

$$\int_0^1|f_x(t)|^p t^{k-1}dt\leqq AB. \tag{64}$$

Moreover

$$\lim_{\epsilon\to 0}\int_0^{\epsilon}|f_x(t)|^p t^{k-1}dt=0.$$

* See S. Bochner, loc. cit. (footnote on p. 179).

The proof follows easily from the relation

$$\int_a^b |f_x(t)|^p t^{k-1}dt \leqq \omega_k \int_S |f(x)|^p dx_1 \cdots dx_k,$$

where ω_k is a number depending only on k, and S is the spherical shell with the radii a and b, and its center at the point x.

LEMMA 5. *Let $f(x); f^1(x), f^2(x), \cdots$ be a sequence of functions each belonging to class L_p over every finite domain, and let*

$$\lim_{n\to\infty} \int_E |f(x) - f^n(x)|^p dx_1 \cdots dx_k = 0 \tag{65}$$

uniformly in all unit spheres E.

Then, uniformly in $\epsilon \leqq t < \infty$ and all x,

$$\lim_{n\to\infty} \int_t^{t+1} |f_x(t) - f_x^n(t)|^p dt = 0, \tag{66}$$

and, for every $\epsilon > 0$, uniformly in $x \equiv (x_1, \cdots, x_k)$,

$$\lim_{n\to\infty} \int_0^\epsilon |f_x(t) - f_x^n(t)|^p t^{k-1}dt = 0. \tag{67}$$

Lemma 5 is an immediate consequence of Lemma 4.

THEOREM IV. *Let $\phi(t)$ be absolutely continuous over every finite range, and let*

$$\int_0^\infty |\phi(t)|\, t^{1+1/2}dt < \infty. \tag{68}$$

Then the formula

$$S_R^\phi(x) = R\int_0^\infty f_x(t)H_\phi(tR)dt \tag{10}$$

holds for every (finite) exponential polynomial

$$\sum a_{n_1\cdots n_k} e^{i(n_1x_1+\cdots+n_kx_k)}$$

the numbers $n_1, \cdots, n_k$ being arbitrary real numbers.

It is sufficient to consider an exponential monomial (21). Formula (10) follows from a combination of (22), (23) with (43)–(45).

THEOREM V. *If $\phi(t)$ is a function as in Lemma 2, then* (10) *holds for every almost periodic function of a Stepanoff class.*

Let

$$f(x) \sim \sum a_{n_1 \cdots n_k} e^{i(n_1 x_1 + \cdots + n_k x_k)} \tag{1}$$

be a function of class S_1, and let

$$f^n(x) = \sum a^{(n)}_{n_1 \cdots n_k} e^{i(n_1 x_1 + \cdots + n_k x_k)} \qquad (n = 1, 2, \cdots) \tag{69}$$

be a sequence of exponential polynomials approximating $f(x)$ in the sense of Lemma 5, with $p = 1$. This approximation implies

$$\lim_{n \to \infty} a^{(n)}_{n_1 \cdots n_k} = a_{n_1 \cdots n_k}. \tag{70}$$

If

$$S^{\phi}_R(x; f^n) = \sum \phi\left(\frac{\nu}{R}\right) a^{(n)}_{n_1 \cdots n_k} e^{i(n_1 x_1 + \cdots + n_k x_k)}, \tag{71}$$

then, by Theorem IV,

$$S^{\phi}_R(x; f^n) = R \int_0^{\infty} f^n_x(t) H_{\phi}(tR) dt. \tag{72}$$

Let now $n \to \infty$ in (72), R being fixed. We split the integral on the right into $\int_0^1$ and $\int_1^{\infty}$. In the first part we may majorize $H_{\phi}(tR)$ by t^{k-1}, in the second part by $t^{-1-\epsilon}$ (see (39), (40)). Applying (67) to the first part and (66) to the second part, we find that the right side of (72) converges towards

$$R \int_0^{\infty} f_x(t) H_{\phi}(tR) dt \tag{73}$$

uniformly in x. Each function $S_R^{\phi}(x; f^n)$ is almost periodic. We have just seen that these functions converge uniformly in x. The limit function is again almost periodic and its Fourier series is the formal limit of the right side of (71). On account of (70) the resulting function has the Fourier series

$$\sum \phi\left(\frac{\nu}{R}\right) a_{n_1 \cdots n_k} e^{i(n_1 x_1 + \cdots + n_k x_k)}. \tag{74}$$

It is therefore well justified to denote the function itself by $S_R^{\phi}(x)$, and the proof of our theorem is completed. We observe that $S_R^{\phi}(x)$ is not only a Stepanoff function but a uniformly continuous almost periodic function in the original sense of H. Bohr.

THEOREM VI. *Theorem* I *holds also for almost periodic functions of the Stepanoff class.*

The proof of Theorem I goes through with the slight change that (52) follows now from (63) instead of from (32).

But if we apply the Riesz kernel $\phi(t) = K_\alpha(t)$ with the critical exponent (12), the series behave differently from the integrals. For $k = 1$, in which case $\alpha = 0$, it is true that the relation (59) depends only on the behaviour of $f(x)$ in a neighborhood of x, at least for all pure periodic functions; this is the content of the well known localization theorem of Riemann. But for every $k \geqq 2$ this is no longer true, even for pure periodic functions, as will appear from the subsequent Theorem VII. Throughout the remainder of this section we shall assume $k \geqq 2$.

We shall denote a general *lattice point* $(m_1, \cdots, m_k)$ by M. We denote these points in some fixed order by $M_1, M_2, M_3, \cdots$. Furthermore we introduce, corresponding to every point M_p, the function

$$\lambda_p(h) = [(m_1 - h_1)^2 + \cdots + (m_k - h_k)^2]^{1/2} \tag{75}$$

of the real variables $h_1, \cdots, h_k$.

LEMMA 6. *For any finite sequence of constants $c_1, \cdots, c_q$ which do not vanish simultaneously, the points $h = (h_1, \cdots, h_k)$ for which*

$$\sum_{p=1}^{q} c_p \lambda_p(h) = 0 \tag{76}$$

are nowhere dense in the Euclidean h-space.

The function $\lambda_p(h)$ is analytic at every point h, with the only exception of the point $h = M_p$ for which this is not the case. One of the numbers c_p is $\neq 0$. Therefore the function on the left side of (76) is certainly singular in at least one point and, consequently, not identically zero. But the zeros of such an analytic function form a (closed) nowhere dense set.

LEMMA 7. *There exists an everywhere dense set of points h for which the sequence of numbers*

$$\lambda_1(h), \lambda_2(h), \cdots \tag{77}$$

is linearly independent with reference to integer (or rational) coefficients.

The desired set is complementary to the set of zeros of all functions $\sum_{p=1}^{q} c_p \lambda_p(h)$ whose coefficients $c_1, \cdots, c_q$ are integers. But the complementary set to an enumerable number of nowhere dense sets is everywhere dense.

LEMMA 8. *If h is not a lattice point and $\beta > 0$, then the sum*

$$C_\beta(\rho) = \sum_{p=1}^{\infty} \frac{e^{2\pi i \rho \lambda_p(h)}}{\lambda_p(h)^{k+\beta}} \tag{78}$$

converges in $-\infty < \rho < \infty$. For every λ the expression

$$\frac{1}{R}\int_1^R C_\beta(\rho)e^{-2\pi i\rho\lambda}d\rho \tag{79}$$

tends to a limit as $R\to\infty$, uniformly in $0<\beta\leqq\beta_0$ (where β_0 is any positive number). This limit has the value

$$\begin{cases} 0 & \text{if } \lambda \neq \lambda_1(h),\ \lambda_2(h),\ \cdots, \\ \lambda_p(h)^{-k-\beta} & \text{if } \lambda = \lambda_p(h). \end{cases} \tag{80}$$

It is immediately seen that $\sum_{p=1}^{\infty}\lambda_p(h)^{-k-\beta}$ is convergent for $\beta>0$, that

$$\sum_{\lambda_p(h)\neq\lambda}\frac{\lambda_p(h)^{-k-\beta}}{\lambda_p(h)-\lambda}$$

converges uniformly in $0<\beta\leqq\beta_0$, and that

$$\left|\frac{1}{R}\int_1^R e^{2\pi i\rho(\lambda_p(h)-\lambda)}d\rho\right| \leqq \frac{2}{R\,|\lambda_p(h)-\lambda|} \quad \text{if } \lambda \neq \lambda_1(h),\ \lambda_2(h),\ \cdots,$$

$$\frac{1}{R}\int_1^R e^{2\pi i\rho(\lambda_p(h)-\lambda)}d\rho = 1-\frac{1}{R} \quad \text{if } \lambda = \lambda_p(h).$$

The assertions of our lemma easily follow if we substitute (78) in (79) and interchange the order of summation and integration.

LEMMA 9. *If h is not a lattice point the function*

$$\Phi_R(h) = \sum_{\nu\leqq R}\left(1-\frac{\nu^2}{R^2}\right)^{\alpha}e^{-2\pi i(n_1h_1+\cdots+n_kh_k)}, \qquad \nu^2 = n_1^2+\cdots+n_k^2, \tag{81}$$

has the following property. For each λ the limit

$$\lim_{R\to\infty}\frac{1}{R}\int_0^R \Phi_\rho(h)e^{-2\pi i\rho\lambda}d\rho$$

exists and equals

$$\begin{cases} \dfrac{\Gamma(\alpha+1)}{2^{1/2}\pi^{k/2}}\,\dfrac{e^{\mp ik\pi/2}}{\lambda_p(h)^k} & \text{if } \lambda = \pm\lambda_p(h), \\ 0 & \text{if } \lambda \neq \pm\lambda_1(h),\ \pm\lambda_2(h),\ \cdots. \end{cases} \tag{82}$$

For $\beta>0$ the function

$$\Phi_R^{\beta}(h) = R^{\beta}\sum_{\nu\leqq R}\left(1-\frac{\nu^2}{R^2}\right)^{\alpha+\beta}e^{-2\pi i(n_1h_1+\cdots+n_kh_k)} \tag{83}$$

has the value

$$\frac{\Gamma(\alpha+1+\beta)}{\pi^{\alpha+\beta}} R^{1/2} \sum_{p=1}^{\infty} \frac{J_{k+\beta-1/2}(2\pi R(\lambda_p(h))}{\lambda_p(h)^{k+\beta-1/2}} \tag{84}$$

provided h is not a lattice point.* For $R \geqq 1$ this is, uniformly in $0 < \beta \leqq \beta_0$,

$$\frac{\Gamma(\alpha+1+\beta)}{2^{1/2}\pi^{k/2+\beta}} \left[e^{-ik\pi/2} C_\beta(R) + e^{ik\pi/2} C_\beta(-R)\right] + O(R^{-1}). \tag{85}$$

Since (81) is the limit of (83) as $\beta \to 0$, Lemma 9 is a consequence of Lemma 8.

LEMMA 10. *If the point h is such that the sequence* (77) *is linearly independent with reference to integer coefficients, then*

$$\limsup_{R\to\infty} \left| \Phi_R(h) \right| = +\infty. \tag{86}$$

The proof follows, on the basis of Lemma 9, from the following lemma by reductio ad absurdum.

LEMMA 11.† *Let $f(x)$ be a real bounded measurable function in $0 < x < \infty$ for which*

$$a(\Lambda) = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(x) e^{-i\Lambda x} dx \tag{87}$$

exists for every Λ. Then $a(\Lambda)$ is different from zero for an at most enumerable set of numbers Λ which we shall denote by $\Lambda_1, \Lambda_2, \cdots$. If the non-negative numbers

$$|\Lambda_1|, |\Lambda_2|, \cdots \tag{88}$$

are linearly independent with reference to integer coefficients, then

$$|a(\Lambda_1)| + |a(\Lambda_2)| + \cdots < \infty. \tag{89}$$

That $a(\Lambda)$ vanishes only for an enumerable number of values follows readily from Bessel's inequality which holds in this case.‡ Denote the numbers (88) by $\mu_1, \mu_2, \cdots$. On putting

$$K(x) = 1 + \frac{e^{ix} + e^{-ix}}{2} = 1 + \cos x \geqq 0,$$

form for any real numbers $\alpha_1, \cdots, \alpha_q$, $q \geqq 2$, the composite kernel

---

* See, for instance, J. R. Wilton, *The lattice points of an n-dimensional ellipsoid*, Journal of the London Mathematical Society, vol. 2 (1927), pp. 227–233; esp. p. 229.

† This lemma was stated without proof in a slightly different wording in S. Bochner, *Über Fourier-Reihen von fastperiodischen Funktionen*, Sitzungsberichte der Berliner Mathematiker Gesellschaft, vol. 26 (1927), p. 54.

‡ See, for instance, M. Fekete, *On generalized Fourier series with the non-negative coefficients*, Proceedings of the London Mathematical Society, vol. 39 (1935), p. 321, footnote.

$$K_q(x) = \prod_{p=1}^{q} K(\mu_p x - \alpha_p)$$

$$= 1 + \frac{1}{2}\sum_{p=1}^{q} (e^{i\mu_p x}e^{-i\alpha_p} + e^{-i\mu_p x}e^{i\alpha_p}) + R_q(x).$$

$R_q(x)$ is an exponential polynomial whose exponents are of the form $\epsilon_1\mu_1+\epsilon_2\mu_2+\cdots+\epsilon_q\mu_q$; $\epsilon_p=0, \pm 1$, and at least two coefficients ϵ_p are different from zero. Owing to our assumption about the linear independence of $\mu_1, \mu_2, \cdots$, the exponents occurring in $R_q(x)$ are all different from $\pm\mu_1, \pm\mu_2, \cdots$. Hence*

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T f(t)K_q(t)dt = \frac{1}{2}\sum_{p=1}^{q} (a(\mu_p)e^{i\alpha_p} + a(-\mu_p)e^{-i\alpha_p}).$$

The absolute value of the left side is not greater than

$$M = \underset{0<x<\infty}{\text{l.u.b.}} |f(x)|.$$

Hence

$$\left|\sum_{p=1}^{q} a(\mu_p)e^{i\alpha_p} + a(-\mu_p)e^{-i\alpha_p}\right| \leqq M. \tag{90}$$

Since $f(x)$ is real, $a(\mu_p)$ and $a(-\mu_p)$ are conjugate complex; for an appropriate choice of α_p

$$a(\mu_p)e^{i\alpha_p} + a(-\mu_p)e^{-i\alpha_p} = |a(\mu_p)| + |a(-\mu_p)|.$$

Since q may be any integer, (90) implies

$$\sum_{p=1}^{\infty} \{|a(\mu_p)| + |a(-\mu_p)|\} < \infty,$$

and this is the same relation as (89).

THEOREM VII. *There exists a function $f(x)=f(x_1, \cdots, x_k)$ of class L having the period 1 in each variable,*

$$f(x) \sim \sum a_{n_1\cdots n_k} e^{2\pi i(n_1x_1+\cdots+n_kx_k)}, \tag{91}$$

and vanishing in a neighborhood of the origin $(0, \cdots, 0)$, such that

$$\limsup_{R\to\infty} \left|\sum_{\rho\leqq R} \left(1 - \frac{\rho^2}{R^2}\right)^{(k-1)/2} a_{n_1\cdots n_k}\right| = \infty. \tag{92}$$

* S. Bochner, *Beiträge zur Theorie der fastperiodischen Funktionen*, I, Mathematische Annalen, vol. 96 (1926), pp. 133–134; H. Bohr, *Fastperiodische Funktionen*, Ergebnisse der Mathematik, Berlin, Springer, 1932, p. 76.

The function of R occurring in (92) can be written as

$$\int_0^1 \cdots \int_0^1 f(x) \sum_{\nu \leqq R} \left(1 - \frac{\nu^2}{R^2}\right)^{(k-1)/2} e^{-2\pi i(n_1 x_1 + \cdots + n_k x_k)} dx_1 \cdots dx_k$$

or, see (81),

$$\int_E f(x) \Phi_R(x) dx_1 \cdots dx_k, \tag{93}$$

where E is a measurable set within the fundamental domain of periodicity outside of which $f(x)$ vanishes.

If our theorem were not true, then for every set E omitting a neighborhood of the origin, (93) would be bounded as $R \to \infty$ for every function $f(x)$ of class L on E. We are now going to refute this possibility. Consider on E the Banach space $\mathfrak{S}$ of functions $f(x)$ belonging to class L, with the usual metric. For each R, (93) defines a functional in $\mathfrak{S}$; we shall denote it by $F_R(f)$. The norm of this functional is

$$\text{ess. l.u.b.}_{x \in E} \, | \Phi_R(x) |.$$

Since $\Phi_R(x)$ is continuous in x, if we assume E to be closed, we can write for this

$$\max_{x \in E} | \Phi_R(x) |.$$

By a known theorem, if a set of functionals $F_R(f)$, is such that for every $f \in \mathfrak{S}$

$$\limsup_{R \to \infty} | F_R(f) | < \infty, \tag{94}$$

then also

$$\limsup_{R \to \infty} \| F_R \| < \infty .*$$

But

$$\limsup_{R \to \infty} \max_{x \in E} | \Phi_R(x) | < \infty$$

implies

$$\limsup_{R \to \infty} | \Phi_R(x) | < \infty, \qquad x \in E. \tag{95}$$

But it follows from Lemmas 7 and 10 that there are sets E for which (95) does not hold in all points x. This proves our theorem.

* See S. Banach, *Théorie des opérations linéaires*, 1932, p. 80, Theorem 5.

We are now going to make a few remarks concerning quadrangular partial sums of the form (24).

LEMMA 11. *If $\phi(t)$ is continuous in $0 \leqq t < \infty$, non-negative, monotonely non-increasing, and if $\int_0^\infty \phi(t)dt < \infty$, then*

$$\lim_{R\to\infty} \frac{1}{R} \sum_{n=0}^{\infty} \phi\left(\frac{a}{R}\, n\right) = \frac{1}{a} \int_0^\infty \phi(t)dt.$$

The proof follows immediately from

$$\int_n^{n+1} \phi\left(\frac{a}{R}\, t\right)dt \leqq \phi\left(\frac{a}{R}\, n\right) \leqq \int_{n-1}^{n} \phi\left(\frac{a}{R}\, t\right)dt, \qquad n \geqq 1.$$

THEOREM VIII. *Let $\phi(t)$ have the following properties. It has two continuous derivatives in $0 \leqq t < \infty$. $\phi(t)$ and $\phi''(t)$ are non-negative and monotonely non-increasing as t increases. Furthermore*

$$\gamma = \int_0^\infty \phi(t)dt < \infty, \qquad \gamma_2 = \int_0^\infty \phi''(t)dt < \infty.$$

Then, for $k \geqq 3$, there exists a function $f(x)$ of period 1 in each variable and vanishing in a neighborhood of the origin, such that

$$\limsup_{R\to\infty} \left| \sigma_R^{\phi}(0) \right| = \infty.$$

Since

$$\sigma_R^{\phi}(0) = \int_0^1 \cdots \int_0^1 f(x) \sum_{\kappa=1}^{k} \Phi_R(x_\kappa)dx_\kappa,$$

where

$$\Phi_R(t) = \sum_{-\infty}^{\infty} \phi\left(\frac{|n|}{R}\right) e^{-2\pi i n t},$$

it is sufficient to show the existence of one point $(x_1^0, \cdots, x_k^0) \not\equiv (0, \cdots, 0) \pmod 1$ for which

$$\limsup_{R\to\infty} \prod_{\kappa=1}^{k} \left| \Phi_R(x_\kappa) \right| = \infty.$$

We choose the point $x_1^0 = \cdots = x_{k-1}^0 = 0$, $x_k^0 = \frac{1}{2}$. Since

$$\Phi_R(0) = \phi(0) + 2\sum_{n=1}^{\infty} \phi\left(\frac{n}{R}\right),$$

therefore, by Lemma 11, $\Phi_R(0)$ has asymptotically, the value γR. On the other hand

$$\Phi_R\left(\frac{1}{2}\right) = \phi(0) - 2\phi\left(\frac{1}{R}\right) + 2\phi\left(\frac{2}{R}\right) - 2\phi\left(\frac{3}{R}\right) + \cdots$$
$$= \sum_{n=0}^{\infty} \int_0^{R^{-1}} \int_0^{R^{-1}} \phi''\left(\frac{2n}{R} + t_1 + t_2\right) dt_1 dt_2.$$

Hence

$$\sum_{n=1}^{\infty} \phi''\left(\frac{2n}{R}\right) \leqq R^2 \Phi_R\left(\frac{1}{2}\right) \leqq \sum_{n=0}^{\infty} \phi''\left(\frac{2n}{R}\right)$$

and therefore, by Lemma 11, $\Phi_R(\frac{1}{2})$ has the asymptotic value $\gamma_2 R^{-1}/2$. Thus, $\Pi_{k=1}\Phi_R(x_\nu^0)$ has, apart from a constant, the asymptotic value R^{k-2} and consequently, is not bounded in R if $k \geqq 3$.

Under Theorem VIII fall such smooth functions as e^{-t}, and $(1-t)^\delta$ for arbitrarily large values of δ. It is true that there are large classes of smooth functions $\phi(t)$ for which the quadrangular sums (24) do obey the localization property. To the latter ones belong e^{-t^2}, and $(1-t^2)^\delta$ for $\delta > k$ (but not for $\delta < k$). But this makes things only worse, in a way. In fact, by a general theorem* summation functions as $(1-t)^\delta$, $(1-t^2)^\delta$ are absolutely equivalent in the case of circular summation, whereas for quadrangular summation this equivalence theorem is apparently lost.

Part IV. Convergence and gap theorems

We shall first give criteria for convergence and absolute convergence of the series (5). They will impose reasonable conditions on the function $f_x(t)$ alone, and no other explicit conditions on the function $f(x_1, \cdots, x_k)$ itself.

Theorem IX. *Let $f(x)$ be a function of class L having the period 2π in each variable. If for a fixed point x the function $f_x(t)$ has k derivatives which are all bounded in $0 \leqq t < \infty$, and if, for k even, the kth derivative of $f_x(t)$ is continuous at $t=0$, then the series* (5) *converges toward $f_x(0) = f(x)$.*

Since our theorem certainly holds for a constant function we may assume $f_x(0) = 0$. For each integer n we put

$$a_n = a_n(x) = \sum_{n=n_1^2+\cdots+n_k^2} a_{n_1 \cdots n_k} e^{i(n_1 x_1 + \cdots + n_k x_k)}. \tag{96}$$

If δ exceeds $(k-1)/2$, then, by Theorem V,

$$\sum_{n \leqq R^2} \left(1 - \frac{n}{R^2}\right)^\delta a_n = R \int_0^\infty f_x(t) H_\delta(tR) dt. \tag{97}$$

* G. H. Hardy, *The second theorem of consistency for summable series*, Proceedings of the London Mathematical Society, vol. 15 (1915), pp. 72–88.

$H_\delta(t)$ has, for some $\kappa>0$, the asymptotic expansion

$$\frac{e^{it}}{t^{1+\kappa}}\sum_{n=0}^{\infty}A_n t^{-n}+\frac{e^{-it}}{t^{1+\kappa}}\sum_{n=0}^{\infty}B_n t^{-n}.$$

Each of the functions

$$P_1(t)=-\int_t^\infty H_\delta(t)dt, \qquad P_2(t)=-\int_t^\infty P_1(t)dt, \cdots$$

has the same property. Hence we may integrate by parts in (97), and the result is:

$$\sum_{n\leqq R^2}\left(1-\frac{n}{R^2}\right)^\delta a_n=\frac{b_1}{R}+\cdots+\frac{b_{k-1}}{R^{k-1}}+\frac{1}{R^k}Q(R), \tag{98}$$

where

$$Q(R)=\int_0^\infty f_x^{(k)}\left(\frac{t}{R}\right)P_k(t)dt.$$

$P_k(t)$ is absolutely integrable over $0\leqq t<\infty$, and $f_x^{(k)}(t)$ is bounded in this interval. Therefore

$$Q(R)=O(1) \qquad \text{as } R\to\infty. \tag{99}$$

If, in addition, $f_x^{(k)}(t)$ is continuous at $t=0$, then

$$Q(R)\to f_x^{(k)}(0)\cdot\int_0^\infty P_k(t)dt \qquad \text{as } R\to\infty. \tag{100}$$

Replacing R^2 by ω in (98) we obtain

$$\sum_{n\leqq\omega}(\omega-n)^\delta a_n=b_1\omega^{\delta-1/2}+b_2\omega^{\delta-1}+\cdots+b_{k-1}\omega^{\delta-(k-1)/2}+\omega^{\delta-k/2}Q(\omega^{1/2}) \tag{101}$$

whereas the assertion of our theorem reads

$$\sum_{n\leqq\omega}a_n=o(1). \tag{102}$$

We shall show that (102) follows from (101) in connection with (100) or (99) respectively.

We first treat the case when k is an even number. The integer $\mu=k/2$ exceeds $(k-1)/2$, thus (101) holds for $\delta=\mu$. Hence

$$\sum_{n\leqq\omega}(\omega-n)^\mu a_n=b_1\omega^{\mu-1/2}+b_2\omega^{\mu-1}+\cdots+b_{k-1}\omega^{1/2}+Q(\omega^{1/2}). \tag{103}$$

In order to deduce (102), we apply to both sides of (103) the operator

$$\begin{aligned}\Delta_{\mu,\eta} A(\omega) &= A(\omega) - \binom{\mu}{1} A(\omega+\eta) + \binom{\mu}{2} A(\omega+2\eta) - \cdots \\ &= (-1)^{\mu} \int_0^{\eta} \cdots \int_0^{\eta} A^{(\mu)}(\omega + t_1 + \cdots + t_{\mu}) dt_1 \cdots dt_{\mu}\end{aligned} \tag{104}$$

for the special values

$$\omega = n + \frac{1}{\mu+2}, \qquad \eta = \frac{1}{\mu+2}. \tag{105}$$

Since

$$\Delta_{\mu,\eta} \sum_{n \leqq \omega} (\omega - n)^{\mu} a_n = (-1)^{\mu} \eta^{\mu} \sum_{n \leqq \omega} a_n \tag{106}$$

and

$$\Delta_{\mu,\eta} (b_1 \omega^{\mu-1/2} + b_2 \omega^{\mu-1} + \cdots) = O(\omega^{-1/2}), \tag{107}$$

we have only to verify that

$$\Delta_{\mu,\eta} Q(\omega^{1/2}) = o(1). \tag{108}$$

But this follows from the relations

$$\Delta_{\mu,\eta} = \Delta_{\mu-1,\eta} (\Delta_{1,\eta}), \tag{109}$$

$$\Delta_{1,\eta} Q(\omega^{1/2}) = o(1), \tag{110}$$

$$\Delta_{\mu-1,\eta} o(1) = o(1), \tag{111}$$

relation (110) being a consequence of (100).

When k is an odd number we put $k = 2\mu+1$. Relation (101) holds for $\delta > \mu$; also, by (99), $Q(\omega^{1/2}) = O(1)$. Hence, for fixed $\epsilon > 0$,

$$\sum_{n \leqq \omega} (\omega - n)^{\mu+\epsilon} a_n = b_1 \omega^{\mu+\epsilon-1/2} + \cdots + O(\omega^{\epsilon-1/2}). \tag{112}$$

We again take, on both sides of (112), the operator (104) for the special values (105), and we obtain

$$\Delta_{\mu,\eta} \sum_{n \leqq \omega} (\omega - n)^{\mu+\epsilon} a_n = O(\omega^{\epsilon-1/2}) + O(\omega^{\epsilon-1/2}). \tag{113}$$

If ϵ is less than $\frac{1}{2}$, then the right side of (113) is $o(1)$; and, introducing the constants

$$\alpha_{\nu} = \int_0^{\eta} \cdots \int_0^{\eta} (\nu + \eta + t_1 + \cdots + t_{\mu})^{\epsilon} dt_1 \cdots dt_{\mu}, \tag{114}$$

we deduce from (113) the set of relations

$$\alpha_0 a_n + \alpha_1 a_{n-1} + \cdots + \alpha_n a_0 = o(1), \qquad n \to \infty .$$

Introducing the partial sums

$$s_n = a_0 + a_1 + \cdots + a_n$$

and the new constants

$$\beta_0 = \alpha_0, \ \beta_1 = \alpha_1 - \alpha_0, \ \beta_2 = \alpha_2 - \alpha_1, \ \cdots , \tag{115}$$

we obtain the new relations

$$\beta_0 s_n + \beta_1 s_{n-1} + \cdots + \beta_n s_0 = o(1),$$

whereas the assertion of our theorem reads

$$s_n = o(1). \tag{116}$$

We consider the analytic function

$$\Phi(z) = \beta_0 + \beta_1 z + \beta_2 z^2 + \cdots$$

and the power series expansion

$$\frac{1}{\Phi(z)} = \gamma_0 + \gamma_1 z + \gamma_2 z^2 + \cdots ,$$

and we shall prove that

$$\sum_{n=0}^{\infty} |\gamma_n| < \infty . \tag{117}$$

Relation (116) is an immediate consequence of (117), since the equations

$$\beta_0 s_n + \beta_1 s_{n-1} + \cdots + \beta_n s_0 = r_n$$

have the unique solution

$$s_n = \gamma_0 r_n + \gamma_1 r_{n-1} + \cdots + \gamma_n r_0 .$$

The proof of (117) will be based on the following two properties of the numbers (115):

$$\beta_n = Bn^{\epsilon-1} + O(n^{\epsilon-2}), \qquad B \neq 0, \tag{118}$$

$$\beta_0 > \beta_1 > \beta_2 > \cdots \geqq 0. \tag{119}$$

In order to secure (119) it is sufficient to fix ϵ so small that $\rho^\epsilon > (\rho+1)^\epsilon - \rho^\epsilon$ for $\rho \geqq \eta$. Both properties can be easily derived from the definition (114).

The function $\Phi(e^{i\theta})$ is continuous and different from zero for $\theta \neq 0$. This follows immediately from

$$(1 - z)\Phi(z) = \beta_0 - (\beta_0 - \beta_1)z - (\beta_1 - \beta_2)z^2 - \cdots$$

in connection with property (119). Property (118) can be rewritten in the form

$$\beta_n = C\binom{n + \epsilon - 1}{n} + O(n^{\epsilon-2}), \qquad C \neq 0;$$

hence

$$\Phi(z) = C(1 - z)^{-\epsilon} + \phi(z),$$

where the function $\phi(z)$ is such that the Fourier series of $\phi(e^{i\theta})$ is absolutely convergent. Since the Fourier series of $(1-e^{i\theta})^{\epsilon}$ is also absolutely convergent, and $\phi(e^{i\theta})\cdot(1-e^{i\theta})^{\epsilon}$ is 0 for $\theta=0$, we conclude that the function

$$\Psi(e^{i\theta}) = (1 - e^{i\theta})^{\epsilon}\Phi(e^{i\theta}) \equiv C + \phi(e^{i\theta})(1 - e^{i\theta})^{\epsilon}$$

is nowhere 0 and has an absolutely convergent Fourier series. By an important theorem of N. Wiener* the Fourier series of $\{\Psi(e^{i\theta})\}^{-1}$ is also absolutely convergent, and so is the Fourier series of $\{\Phi(e^{i\theta})\}^{-1} = (1-e^{i\theta})^{\epsilon}\cdot\{\psi(e^{i\theta})\}^{-1}$. But the numbers γ_n are the Fourier coefficients of the latter function and thus (117) is proved.

THEOREM X. *Let $f(x)$ be a function of class L having the period 2π in each variable. If for a fixed point x the function $f_x(t)$ has $k+4$ derivatives which are all bounded in $0 \leqq t < \infty$, then the simple series* (5) *converges absolutely toward $f_x(0) = f(x)$.*

The theorem asserts that, under the stated assumptions,

$$\sum_{n=0}^{\infty} |\, a_n(x)\,| < \infty\,. \tag{120}$$

Replacing k by $k+4$ in the right side of (98) we obtain the relation

$$\sum_{n \leqq \omega} (\omega - n)^{\delta} a_n = b_1\omega^{\delta-1/2} + b_2\omega^{\delta-1} + \cdots + O(\omega^{\delta-(k+4)/2}). \tag{121}$$

Let μ denote the smallest integer exceeding $(k-1)/2$, thus $\mu \leqq (k+1)/2$. Putting $\delta = \mu$ in (121) and applying the operator (104) for the values (105) we obtain

$$\sum_{n \leqq \omega} a_n = c_1\omega^{-1/2} + c_2\omega^{-1} + O(\omega^{-3/2}) + O(\omega^{(k+1)/2-(k+4)/2}).$$

Hence

* N. Wiener, *Tauberian theorems*, Annals of Mathematics, (2), vol. 33 (1932), p. 14. Compare also R. E. A. C. Paley and N. Wiener, *Fourier Transforms in the Complex Domain*, American Mathematical Society Colloquium Publications, vol. 19, New York, 1934, p. 59.

$$a_n = a_n(x) = O(n^{-3/2}).$$

THEOREM XI. *Let $f(x)$ be an almost periodic function of class S whose Fourier exponents $(n_1, \cdots, n_k)$ have no finite accumulation point, and let its Fourier series be written in the form*

$$\sum_{n=0}^{\infty} a_{\lambda_n}(x) \tag{122}$$

$$\lambda_1 < \lambda_2 < \lambda_3 < \cdots \to \infty,$$

where

$$a_\lambda(x) = \sum_{\lambda = n_1^2 + \cdots + n_k^2} a_{n_1 \cdots n_k} e^{i(n_1 x_1 + \cdots + n_k x_k)}.$$

Suppose further that for a subsequence $\{\lambda_{n_j}\}$ of $\{\lambda_n\}$ and a constant $\sigma \geqq 0$,

$$(\lambda_{n_j+1} - \lambda_{n_j})^{-1} = O(\lambda_{n_j}^{\sigma}). \tag{123}$$

If, for a fixed point x, $f_x(t)$ has more than $(k+1)(1+\sigma)$ derivatives each of which is bounded in $0 \leqq t < \infty$, then

$$\lim_{j \to \infty} \sum_{n=0}^{n_j} a_{\lambda_n}(x) = f(x).$$

We again assume $f(x) = 0$. If $f_x(t)$ has q derivatives, then, in analogy to (121),

$$\sum_{\lambda_n \leqq \omega} (\omega - \lambda_n)^{\delta} a_{\lambda_n}(x) = b_1 \omega^{\delta - 1/2} + \cdots + O(\omega^{\delta - q/2}). \tag{124}$$

Let $a > 0$ be so chosen that for all large values of j

$$\lambda_{n_j+1} - \lambda_{n_j} \geqq \mu a \lambda_{n_j}^{-\sigma},$$

where μ is again the least integer exceeding $(k-1)/2$. We apply to (124), with $\delta = \mu$, the operator (104) for the values

$$\omega = \lambda_{n_j}, \qquad \eta = a\lambda_{n_j}^{-\sigma} = a\omega^{-\sigma},$$

and we obtain

$$\eta^{\mu} \sum_{\lambda_n \leqq \omega} a_{\lambda_n}(x) = O(\eta^{\mu} \omega^{-1/2}) + O(\omega^{\mu - q/2});$$

hence

$$\sum_{\lambda_n \leqq \omega} a_{\lambda_n}(x) = o(1) + O(\omega^{\mu + \mu\sigma - q/2}).$$

But for $q > (k+1)(1+\sigma)$ the last term is $o(1)$, and the proof of our theorem is completed.

Theorems IX–XI give criteria for Riesz summability of order $\delta=0$. The following theorems will be concerned with Riesz summability for the critical exponent $\delta=\alpha$.

THEOREM XII. *Let $f(x)$ be an almost periodic function of class S_p, $p>1$, whose Fourier exponents have no finite accumulation point. If for a given point x, $f_x(t)$ satisfies the (local) Lebesgue condition* (61) *then, for every $q>0$, no matter how large,*

$$\lim_{R\to\infty}\int_R^{R+1}\left|S_\rho^\alpha(x)-2^l\Gamma(l+1)\cdot s\right|^q d\rho=0. \tag{125}$$

Since the theorem is obviously true for a constant function we can assume $s=0$. Since the class S_p is the more general the smaller p is and relation (125) is the more general the larger q is, we may assume in the proof of the theorem that

$$q=\frac{p}{p-1}. \tag{126}$$

Let $f^n(x)$ be a sequence of exponential polynomials as in Lemma 5. They may be so chosen that no new Fourier exponents $(n_1,\cdots,n_k)$ arise. Since the Fourier exponents of the given function $f(x)$ have no finite point of accumulation, it is trivial that $S_R(x,f^n)$ (see (71)) tends to $S_R(x,f)$ as $n\to\infty$ for every $\delta\geqq 0$, in particular for $\delta=\alpha$.

By Lemmas 4 and 5,

$$\lim_{n\to\infty}\int_0^\epsilon\left|f_x(t)-f_x^n(t)\right|^p t^{k-1}dt=0, \tag{127}$$

$$\int_\epsilon^\infty\left|\frac{f_x(t)}{t}\right|^p dt<\infty, \tag{128}$$

$$\lim_{n\to\infty}\int_\epsilon^\infty\left|\frac{f_x(t)}{t}-\frac{f_x^n(t)}{t}\right|^p dt=0; \tag{129}$$

and the condition (61) reads for $s=0$,

$$\frac{1}{\eta^k}\int_0^\eta t^{k-1}\left|f_x(t)\right|dt\to 0,\qquad \int_\eta^\epsilon\frac{\left|f_x(t)-f_x(t+\eta)\right|}{t}dt\to 0 \text{ as } \eta\to 0. \tag{130}$$

By Theorem IV

$$S_R^\alpha(x;f^n)=R\int_0^\infty f_x^n(t)H_\alpha(tR)dt=P_R^\alpha(x,f^n)+Q_R^\alpha(x,f^n), \tag{131}$$

where

$$P_R^\alpha(x, f) = R\int_0^\epsilon f_x(t)H_\alpha(tR)dt$$

$$Q_R^\alpha(x, f) = R\int_\epsilon^\infty f_x(t)H_\alpha(tR)dt. \tag{132}$$

Since, for fixed R, $H_\alpha(tR)$ is majorized by t^{k-1} in $0 \leqq t \leqq \epsilon$, it follows from (127), if we use it for $p=1$,

$$\lim_{n\to\infty} P_R^\alpha(x, f^n) = P_R^\alpha(x, f). \tag{133}$$

Furthermore, using the asymptotic expansion

$$RH_\alpha(tR) \sim \frac{e^{itR}}{t}\sum_0^\infty A_n(tR)^{-n} + \frac{e^{-itR}}{t}\sum_0^\infty B_n(tR)^{-n} \tag{134}$$

which is valid for $tR\to\infty$, the argument leading to Theorem III shows that, under condition (130), $P_R^\alpha(x, f)\to 0$ as $R\to\infty$. In particular,

$$\lim_{R\to\infty}\int_R^{R+1} |P_R^\alpha(x, f)|^q dR = 0. \tag{135}$$

The term $Q_R^\alpha(x, f)$ is less simple. Using (128) and (134) we see as in the theory of Hankel transforms,* that $Q_R^\alpha(x, f)$ exists if the integral (132) is defined as a limit-in-mean, and that

$$\int_1^R |Q_R^\alpha(x, f)|^q dR < \infty.$$

In particular

$$\lim_{R\to\infty}\int_R^{R+1} |Q_R^\alpha(x, f)|^q dR = 0. \tag{136}$$

Furthermore, on account of (129),

$$\lim_{n\to\infty}\int_1^\infty |Q_R^\alpha(x, f^n) - Q_R^\alpha(x, f)|^q dR = 0.$$

Hence, for almost all R,

$$S_R^\alpha(x, f) = P_R^\alpha(x, f) + Q_R^\alpha(x, f). \tag{137}$$

* See E. C. Titchmarsh, *A note on Hankel transforms*, Journal of the London Mathematical Society, vol. 1 (1926), pp. 195–196.

Relation (125) follows now from (135)–(137).

LEMMA 12. *Let $f(x)$ be an almost periodic function of class S whose Fourier exponents have no finite point of accumulation. Then the relation*

$$S_R^a(x, f) = R\int_0^\infty f_x(t)H_a(tR)dt \tag{138}$$

holds in the following integrated form. If $A(R)$ is any absolutely continuous function in any finite interval $(0<)R_1 \leqq R \leqq R_2$, then

$$\int_{R_1}^{R_2} A(R)S_R^a(x)dR = \int_0^\infty f_x(t)dt \int_{R_1}^{R_2} A(R)RH_a(tR)dR. \tag{139}$$

We again introduce approximating exponential polynomials $f^n(x)$. The relation (138) holds for $f_x^n(t)$, and since, for each n,

$$\int_0^\infty |f_x^n(t)|\,|H_a(tR)|\,dt < \infty$$

uniformly in $R_1 \leqq R \leqq R_2$, (139) is true for $f^n(x)$. Thus we have only to show that

$$\lim \int_0^\infty f_x^n(t)dt \int_{R_1}^{R_2} A(R)RH_a(tR)dR = \int_0^\infty f_x(t)dt \int_{R_1}^{R_2} A(R)RH_a(tR)dR. \tag{140}$$

Now, using (134), it is easily seen that

$$\int_{R_1}^{R_2} A(R)RH_a(tR)dR = O(t^{-2}) \quad \text{as } t \to \infty; \tag{141}$$

on the other hand

$$\int_{R_1}^{R_2} A(R)RH_a(tR)dR = O(t^{k-1}) \text{ as } t \to 0. \tag{142}$$

Relation (140) follows now from a combination of (141), (142) with Lemma 5.

LEMMA 13. *Let $f(x)$ be an almost periodic function of class S whose Fourier exponents have no finite point of accumulation. If, for a given point x, $f_x(t)$ satisfies the (local) Lebesgue condition (61), then, for any $\gamma \geqq 0$, and any function $P(\rho)$ which has a bounded derivative in $-1 \leqq \rho \leqq 1$,*

$$\lim_{R\to\infty} \int_{-1}^{1} P(\rho)\left(1+\frac{\rho}{R}\right)^\gamma S_{R+\rho}^a(x)d\rho = 2^l\Gamma(l+1)\cdot s\int_{-1}^{1} P(\rho)d\rho.$$

We may again assume $s=0$. By Lemma 12, for any $a>0$, the expression on the left may be written as the sum of the terms

$$\int_{-1}^{1} d\rho P(\rho)\left(1+\frac{\rho}{R}\right)^{\gamma}\cdot(R+\rho)\int_{0}^{a} f_x(t)H_\alpha(t(R+\rho))dt \tag{143}$$

$$\int_{a}^{\infty} dt f_x(t)\int_{-1}^{1} P(\rho)\left(1+\frac{\rho}{R}\right)^{\gamma}(R+\rho)H_\alpha(t(R+\rho))d\rho. \tag{144}$$

The argument leading up to Theorem III shows that under the condition (61), for every $a>0$,

$$\lim_{R\to\infty} R\int_{0}^{a} f_x(t)H_\alpha(tR)dt = 0. \tag{145}$$

Hence (143) goes to 0 if $R\to\infty$; and we have still to prove that (144) will be sufficiently small, uniformly in $R\geqq 2$, if only a is sufficiently large. The latter will obviously be the case if the inner integral in (144) is $O(t^{-2})$ for $t\to\infty$, uniformly in $R\geqq 2$. But, making use of (134), we see that the inner integral in (144) is, uniformly in $R\geqq 2$,

$$\begin{aligned} &A_0\int_{-1}^{1} P(\rho)\left(1+\frac{\rho}{R}\right)^{\gamma}\frac{e^{it(R+\rho)}}{t}d\rho \\ &\qquad + B_0\int_{-1}^{1} P(\rho)\left(1+\frac{\rho}{R}\right)^{\gamma}\frac{e^{-it(R+\rho)}}{t}d\rho + O(t^{-2}). \end{aligned} \tag{146}$$

But the relation

$$\begin{aligned} \int_{-1}^{1} P(\rho)\left(1+\frac{\rho}{R}\right)^{\gamma}e^{it\rho}d\rho &= \frac{1}{it}\left[P(\rho)\left(1+\frac{\rho}{R}\right)^{\gamma}e^{it\rho}\right]_{-1}^{1} \\ &\quad - \frac{1}{it}\int_{-1}^{1}\frac{d}{d\rho}\left[P(\rho)\left(1+\frac{\rho}{R}\right)^{\gamma}\right]e^{it\rho}d\rho \end{aligned} \tag{147}$$

shows that (146) is also $O(t^{-2})$, uniformly in $R\geqq 2$, and the proof of the lemma is completed.

LEMMA 14. *Given a positive integer g, there exists a polynomial $P(\rho)=A_0+A_1\rho+A_2\rho^2+\cdots+A_g\rho^g$ such that*

$$\int_{-1}^{1} P(\rho)d\rho = 1, \tag{148}$$

$$\int_{-1}^{1} P(\rho)\rho^{\lambda}d\rho = 0, \qquad \lambda = 1, 2, \cdots, g. \tag{149}$$

In order to obtain such a polynomial we may apply E. Schmidt's process of orthogonalization to the functions $t, t^2, \cdots, t^g, 1$ (in this order) on the interval $(-1, 1)$.

LEMMA 15. *Assumptions*: (i) $P(t)$ *satisfies* (148), (149) *for some* $g>\alpha$; (ii) $N_1, N_2, \cdots$ *is a sequence of points in the k-dimensional space having no finite point of accumulation and* ν_p *denotes the distance of* N_p *from the origin*; (iii) *the constants* $a(N_1), a(N_2), \cdots$ *are such that*

$$\sum_{R\leqq\nu_p\leqq R+1} |\, a(N_p)\,| = o(R^\alpha) \quad \text{as } R\to\infty;$$

(iv) $A(R)$ *and* $B(R)$ *are defined by*

$$A(R) = \sum_{\nu_p\leqq R} (R^2-\nu_p^2)^\alpha a(N_p), \tag{150}$$

$$B(R) = \int_{-1}^{1} A(R+\rho)P(\rho)d\rho. \tag{151}$$

Conclusion:

$$\lim_{R\to\infty} \frac{A(R)-B(R)}{R^{2\alpha}} = 0. \tag{152}$$

It follows from (iii) that

$$\sum_{R-2\leqq\nu_p\leqq R} (R^2-\nu_p^2)^\alpha a(N_p) = o(R^\alpha)\left[R^2-(R-2)^2\right]^\alpha = o(R^{2\alpha}). \tag{153}$$

Hence

$$A(R) - \sum_{\nu_p\leqq R-2} = o(R^{2\alpha}). \tag{154}$$

Substituting this in the right side of $B(R)$ we easily find

$$B(R) - \sum_{\nu_p\leqq R-2} \int_{-1}^{1} ((R+\rho)^2-\nu_p^2)^\alpha P(\rho)d\rho a(N_p) = o(R^{2\alpha}). \tag{155}$$

Hence (152) will be proved if we show

$$\sum_{\nu_p\leqq R-2} C(R,\nu_p)a(N_p) = o(R^{2\alpha}), \tag{156}$$

where

$$C(R,\nu) = (R^2-\nu^2)^\alpha - \int_{-1}^{1} ((R+\rho)^2-\nu^2)^\alpha P(\rho)d\rho. \tag{157}$$

Now

$$\begin{aligned}((R+\rho)^2-\nu^2)^\alpha &= (R^2-\nu^2+2R\rho+\rho^2)^\alpha \\ &= (R^2-\nu^2)^\alpha(1+b_1\rho+\cdots+b_g\rho^g+b_{g+1}(\rho)\rho^{g+1}),\end{aligned}$$

where $b_{g+1}(\rho)$ uniformly in $\nu\leqq R-2$ and $|\rho|\leqq 1$, is majorized by

$$\left| \frac{2R}{R^2 - \nu^2} + \frac{1}{R^2 - \nu^2} \right|^{\rho+1}$$

which in its turn is majorized by

$$\frac{1}{(R - \nu)^{\rho+1}}.$$

Applying the decisive properties (148), (149) of $P(\rho)$ we find that, uniformly in $\nu \leqq R-2$, $C(R, \nu)$ is majorized by

$$\frac{R^\alpha}{(R - \nu)^{\rho+1-\alpha}}.$$

Moreover, for every fixed ν, $C(R, \nu) = o(R^{2\alpha})$. Therefore, the left side of (156) is

$$o(1) \cdot R^\alpha \int_0^{R-2} \frac{\nu^\alpha d\nu}{(R - \nu)^{\rho+1-\alpha}}. \tag{158}$$

But for $\rho+1-\alpha>1$, that is, $\rho>\alpha$, the factor of $o(1)$ in (156) is $o(R^{2\alpha})$ and this proves relation (156).

THEOREM XIII. *Let $f(x)$ be an almost periodic function of class S whose Fourier exponents have no finite point of accumulation. If its Fourier coefficients are such that*

$$\sum_{R \leqq \nu_p \leqq R+1} |a_{n_1 \cdots n_k}| = o(R^{(k-1)/2}), \tag{159}$$

then the (local) Lebesgue condition (61) *is sufficient for the validity of* (59).

The proof follows from Lemma 15, and Lemma 13 with $\gamma = 2\alpha$.

The number of lattice points in the k-dimensional space for which $R \leqq \nu_p \leqq R+1$ is $O(R^{k-1})$ which, for $k \geqq 2$, is larger than $O(R^{(k-1)/2})$. Hence, for a periodic function of class L, (159) need not hold if $k \geqq 2$, and Theorem VII shows that (159) cannot be so relaxed as to be satisfied by any periodic function. But it follows from Bessel's inequality that (159) does hold if $f(x)$ belongs to the class L_2.

COROLLARY TO THEOREM XIII. *If $f(x)$ is a periodic function belonging to class L_2, then the (local) Lebesgue condition* (61) *is sufficient for the validity of* (59).

It would be interesting to decide if this also holds for periodic functions belonging to class L_p, for $1<p<2$.

PRINCETON UNIVERSITY,
PRINCETON, N. J.

ON CERTAIN GENERALIZATIONS OF THE CAUCHY-TAYLOR EXPANSION THEORY*

BY

GERTRUDE STITH KETCHUM

I. Introduction

The purpose of this paper is to consider the expression of any function $f(x)$ analytic at zero in an expansion of the form

$$\sum_{n=0}^{\infty} \alpha_n F_n(x) \tag{1}$$

where the functions $F_n(x)$ are given by

$$F_n(x) = x^n + \sum_{s=1}^{\infty} a_s^{(n)} x^{n+s}, \tag{1.1}$$

$F_n(x)$ are analytic in the region $|x| < r$, and $|F_n(x)/x^n|$ are uniformly bounded in the same region.

Pincherle† as early as 1881 showed that such expansions are always possible. We therefore are interested not in the possibility of such expansions but rather in the region of their validity. The facts which have been obtained for this problem from studies of expansions in functions of more general type seem to be of a restricted nature.‡ Okada§ published an important paper on this subject in 1922. Since that time the problem has been studied by numerous other writers.‖ The work of Pincherle and Okada has generally been overlooked by later writers. However, the importance of the work of Okada is shown by the fact that it includes as special cases theorems obtained by three of the later writers, namely, Izumi, Widder, and Takahashi.¶ The proofs

* Presented to the Society, April 20, 1935; received by the editors September 23, 1935.

† Memorie della Academia delle Scienze dell' Istituto di Bologna, ser. 4, vol. 3 (1881), pp. 151 ff. See also L. Tonelli, Annali di Matematica, ser. 3, vol. 18, pp. 117 ff.

‡ See, for instance, G. D. Birkhoff, Comptes Rendus, vol. 164 (1917), pp. 942 ff.

§ Tôhoku Mathematical Journal, vol. 22 (1922–23), pp. 325 ff.

‖ S. Izumi, Tôhoku Mathematical Journal, vol. 28 (1927), pp. 97 ff. R. F. Graesser, American Journal of Mathematics, vol. 49 (1927), pp. 577–597. S. Narumi, Tôhoku Mathematical Journal, vol. 30 (1928–29), pp. 441 ff. D. V. Widder, these Transactions, vol. 31 (1929), pp. 43–53. S. Takenaka, Proceedings of the Tokyo Physico-Mathematical Society, ser. 3, vol. 13 (1931), pp. 111–117. S. Takahashi, Tôhoku Mathematical Journal, vol. 35 (1932), pp. 242 ff.

¶ A theorem of a different form which is also an easy consequence of Okada's theorem is due to Carmichael. See Bulletin of the American Mathematical Society, vol. 40 (1934), p. 211 (abstract). Theorem VI is a generalization of this theorem due to Carmichael.

offered for these special cases are essentially at least as difficult as the proof of Okada's theorem. Moreover, we give (Theorem III) a slight generalization of his work which includes all the general expansion theorems *on this problem* in the papers cited.

Theorem III, however, is not sufficient to yield the facts obtained by Graesser* concerning expansions in a particular set of confluent hypergeometric functions. Further we show (§V) that these facts cannot be obtained from any theorem, such as III, in which each condition is stated in terms of any set of positive bounds on the coefficients $a_s^{(n)}$. Hence, instead of using bounds on a single coefficient $a_s^{(n)}$, we consider bounds on certain linear combinations of these coefficients, and thereby introduce an auxiliary set of multipliers. We then obtain a general expansion principle (Theorem I) from which by a suitable choice of the multipliers we derive all the known expansion theorems as well as new theorems of the type considered. In particular, by choosing the multipliers from certain binomial sums, we obtain (§V) a generalization of the expansion in hypergeometric functions of Graesser.

The value of Theorem I in actual application rests on a good choice of certain sets of arbitrary constants. Theorems II and III (§II) are simpler and more readily applied. With the exception of the work of Okada and Graesser mentioned above, the known theorems on this subject and generalizations of these theorems are derived in §III. These theorems readily adapt themselves to applications.

In §IV, we obtain a class of sets of integral functions $F_n(x)$ having the property that the expansion of a function $f(x)$ analytic at zero will always converge absolutely and uniformly in any region interior to the circle through a singularity of $f(x)$ nearest the origin. Moreover, this class of functions includes the Bessel coefficients. Expansions in restricted sets of integral functions $F_n(x)$ with leading term x^n have also been treated (see §III) by Julia,† Onofri,‡ and Valiron.§

We apply the results of §III to obtain (§VI) new general theorems on expansions in products of functions of the form $F_n(x)$. These theorems are in turn applied to obtain very considerable generalizations of the known expansions in products of Bessel functions.

II. General theorems

Consider a set of functions $F_n(x)$ defined as in (1.1) and analytic in the

* Loc. cit., p. 592.

† G. Julia, Acta Mathematica, vol. 54 (1930), pp. 280 ff.

‡ L. Onofri, Annali di Matematica, ser. 4, vol. 13 (1934–35), pp. 209 ff.

§ G. Valiron, Bulletin des Sciences Mathématiques, vol. 69 (1934), pp. 26–28.

region $|x|<r$. As indicated in the introduction, we consider positive numbers $b_\nu^{(\mu)}$ which are bounds for linear combinations of the coefficients $a_\nu^{(\mu)}$, thus

$$b_\nu^{(\mu)} \geqq \left| \sum_{t=0}^{\nu} k_{\mu+\nu}^{(\mu+t)} a_t^{(\mu)} \right|,$$

where $a_0^{(\mu)}=1$. Here the multipliers $k_{\mu+\nu}^{(\mu)}$ are complex numbers at one's choice subject to the condition that $k_\mu^{(\mu)}=1$. Let $\delta_{\mu+\nu}^{(\mu)}$ be positive numbers such that $\delta_{\mu+\nu}^{(\mu)} \geqq |k_{\mu+\nu}^{(\mu)}|$. Put

$$M_{\mu+\nu}^{(\mu)} = b_1^{(\mu+\nu)} + \max\left(\delta_{\mu+\nu+1}^{(\mu)}/\delta_{\mu+\nu}^{(\mu)}, b_{\nu+1}^{(\mu)}/b_\nu^{(\mu)}, b_\nu^{(\mu+1)}/b_{\nu-1}^{(\mu+1)}, \cdots, b_2^{(\mu+\nu-1)}/b_1^{(\mu+\nu-1)}\right).$$

Denote by $\overline{M}_{\mu+\nu}^{(\mu)}$ expressions obtained from $M_{\mu+\nu}^{(\mu)}$ by omitting any or all of the outstanding b's or of ratios of b's with a superscript μ_1 such that

$$\sum_{t=0}^{p} k_{\mu_1+p}^{(\mu_1+t)} a_t^{(\mu_1)} = 0 \text{ for all } p>0.$$

Assuming the notation and conditions stated above we have

THEOREM I. *If there exist numbers $k_{\mu+\nu}^{(\mu)}$, $b_\nu^{(\mu)}$, $\delta_{\mu+\nu}^{(\mu)}$, and a positive number R such that*

$$\left|\frac{F_n(x)}{x^n}\right| \leqq Q \text{ for } |x| \leqq R < r, \tag{2.1}$$

where Q is independent of n, and such that

$$\sum_{\mu,\nu=0}^{\infty} |\gamma_\mu| \overline{M}_{\mu+\nu-1}^{(\mu)} \overline{M}_{\mu+\nu-2}^{(\mu)} \cdots \overline{M}_\mu^{(\mu)} \delta_\mu^{(\mu)} R^{\mu+\nu} \tag{2.2}$$

converges, then the function $f(x)\equiv\sum_{\mu=0}^{\infty}\gamma_\mu x^\mu$ has an absolutely and uniformly convergent expansion of the form (1) valid for $|x|\leqq R$. Further, there is only one such expansion converging uniformly for $|x|\leqq R'$, $R'>0$.*

The proof will be given for the particular case $\overline{M}_{\mu+\nu}^{(\mu)}=M_{\mu+\nu}^{(\mu)}$ but the necessary changes to obtain the more general case are obvious.

Let $c_{n+\nu}^{(n)}$, $n=0, 1, 2, \cdots$, be defined by the equations†

$$\begin{aligned} &c_n^{(n)} = 1, \\ &c_{n+\nu}^{(n)} + c_{n+\nu-1}^{(n)} a_1^{(n+\nu-1)} + c_{n+\nu-2}^{(n)} a_2^{(n+\nu-2)} + \cdots + c_n^{(n)} a_\nu^{(n)} = 0, \quad \nu>0. \end{aligned} \tag{2.3}$$

* We shall see later that if (2.1) is satisfied for a value $R_1>0$ and $f(x)$ is analytic for $|x|\leqq R_1$, it is possible to choose the quantities $k_{\mu+\nu}^{(\mu)}$, $\delta_{\mu+\nu}^{(\mu)}$, $b_\nu^{(\mu)}$ so that (2.2) converges for some positive $R\leqq R_1$. In applications the problem is to choose these quantities so as to obtain a value of R as large as possible.

† These equations are obtained by equating coefficients of x^ν in the formal relation

$$x^n=\sum_{p=0}^{\infty} c_{n+p}^{(n)} F_{n+p}(x).$$

Then

$$(2.4)\qquad c_{n+\nu}^{(n)} = (-1)^{\nu}\begin{vmatrix} 1 & 1 & 0 & 0 & \cdots 0 \\ 0 & a_1^{(n)} & 1 & 0 & \cdots 0 \\ 0 & a_2^{(n)} & a_1^{(n+1)} & 1 & \cdots 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots \\ 0 & a_{\nu-1}^{(n)} & a_{\nu-2}^{(n+1)} & a_{\nu-3}^{(n+2)} & \cdots 1 \\ 0 & a_{\nu}^{(n)} & a_{\nu-1}^{(n+1)} & a_{\nu-2}^{(n+2)} & \cdots a_1^{(n+\nu-1)} \end{vmatrix}.$$

We next modify the determinant in (2.4) by multiplying the rth row, $r=1, 2, \cdots, p-1$, by $k_{n+p-1}^{(n+r-1)}$ and adding to the pth row. Hence

$$|c_{n+\nu}^{(n)}| \leqq \left\{ \begin{array}{llllll} \delta_n^{(n)} & 1 & 0 & 0 & \cdots 0 \\ \delta_{n+1}^{(n)} & b_1^{(n)} & 1 & 0 & \cdots 0 \\ \delta_{n+2}^{(n)} & b_2^{(n)} & b_1^{(n+1)} & 1 & \cdots 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots \\ \delta_{n+\nu-1}^{(n)} & b_{\nu-1}^{(n)} & b_{\nu-2}^{(n+1)} & b_{\nu-3}^{(n+2)} & \cdots 1 \\ \delta_{n+\nu}^{(n)} & b_{\nu}^{(n)} & b_{\nu-1}^{(n+1)} & b_{\nu-2}^{(n+2)} & \cdots b_1^{(n+\nu-1)} \end{array} \right\} \equiv A_{n+\nu}^{(n)},$$

where the expression in braces is to be expanded by the same law as a determinant except that all signs are taken positive.

Hence expanding $A_{n+\nu}^{(n)}$ by the last column

$$|c_{n+\nu}^{(n)}| \leqq b_1^{(n+\nu-1)} A_{n+\nu-1}^{(n)} + \left\{ \begin{array}{llllll} \delta_n^{(n)} & 1 & 0 & \cdots 0 \\ \delta_{n+1}^{(n)} & b_1^{(n)} & 1 & \cdots 0 \\ \delta_{n+2}^{(n)} & b_2^{(n)} & b_1^{(n+1)} & \cdots 0 \\ \cdot & \cdot & \cdot & \cdots \\ \delta_{n+\nu-2}^{(n)} & b_{\nu-2}^{(n)} & b_{\nu-3}^{(n+1)} & \cdots 1 \\ \delta_{n+\nu}^{(n)} & b_{\nu}^{(n)} & b_{\nu-1}^{(n+1)} & \cdots b_2^{(n+\nu-2)} \end{array} \right\},$$

$$\begin{aligned} &\leqq b_1^{(n+\nu-1)} A_{n+\nu-1}^{(n)} + A_{n+\nu-1}^{(n)} \\ &\qquad \cdot \max\left(\delta_{n+\nu}^{(n)}/\delta_{n+\nu-1}^{(n)},\ b_{\nu}^{(n)}/b_{\nu-1}^{(n)},\ b_{\nu-1}^{(n+1)}/b_{\nu-2}^{(n+1)}, \cdots, b_2^{(n+\nu-2)}/b_1^{(n+\nu-2)}\right), \\ &\leqq M_{n+\nu-1}^{(n)} A_{n+\nu-1}^{(n)}, \\ &\leqq M_{n+\nu-1}^{(n)} M_{n+\nu-2}^{(n)} \cdots M_n^{(n)} \delta_n^{(n)}. \end{aligned}$$

Thus the double series $\sum_{\mu,\nu=0}^{\infty} \gamma_\mu c_{\mu+\nu}^{(\mu)} F_{\mu+\nu}(x)$ is dominated for $|x| \leqq R$ by the series (2.2) multiplied by Q. Hence the series

$$\sum_{\mu=0}^{\infty}\Big(\sum_{\nu=0}^{\mu}\gamma_{\mu-\nu}c_{\mu}^{(\mu-\nu)}\Big)F_{\mu}(x)$$

converges absolutely and uniformly for $|x|\leqq R$ to an analytic function $g(x)$. By a theorem due to Weierstrass the coefficients of powers of x in the expansion of $g(x)$ may be obtained by expanding $F_{\mu}(x)$ and collecting like powers of x. But it is readily seen from the relations (2.3) that the coefficient of x^n is γ_n. Hence $g(x)=f(x)$.

THEOREM II. *Let the functions $F_n(x)$ be defined as in* (1.1) *and analytic in the region $|x|<r$. Let $b_{\nu}^{(\mu)}$ be any set of positive numbers such that $b_{\nu}^{(\mu)}\geqq|a_{\nu}^{(\mu)}+k_{\mu+\nu}a_{\nu-1}^{(\mu)}|$, $\nu=1, 2, \cdots$; $\mu=0, 1, \cdots$, where $a_0^{(\mu)}=1$ and k_n is a set of complex numbers such that* $\limsup_{n\to\infty}|k_n|^{1/n}\leqq 1$. *If $R<\rho\leqq r$ and*

$$0<R<\liminf_{n\to\infty}M_n^{-1}, \tag{2.5}$$

$$M_n\equiv b_1^{(n)}+\max\{b_{n+1}^{(0)}/b_n^{(0)}, b_n^{(1)}/b_{n-1}^{(1)}, \cdots, b_2^{(n-1)}/b_1^{(n-1)}\}$$
$$(n=0, 1, 2, \cdots),$$

and if

$$\left|\frac{F_n(x)}{x^n}\right|\leqq Q \text{ for } |x|=R, \tag{2.6}$$

where Q is independent of n, then any function $f(x)$ analytic for $|x|<\rho$ has an absolutely and uniformly convergent expansion of the form (1) *valid for $|x|\leqq R$. Further, this is the only such expansion converging uniformly for $|x|\leqq R'$, $R'>0$.*

In Theorem I take $k_{\mu+1}^{(\mu)}=k_{\mu+1}$, $k_{\mu+p}^{(\mu)}=0$, $p=2, 3, \cdots$. For every positive ϵ there corresponds a K_{ϵ} greater than ϵ and such that $|k_{\mu+1}^{(\mu)}|\leqq K_{\epsilon}(1+\epsilon)^{\mu}$. Now let $\delta_{\mu+\nu}^{(\mu)}=K_{\epsilon}\epsilon^{\nu-1}(1+\epsilon)^{\mu}$. Then

$$M_{\mu+\nu}^{(\mu)}=b_1^{(\mu+\nu)}+\max(\epsilon, b_{\nu+1}^{(\mu)}/b_{\nu}^{(\mu)}, b_{\nu}^{(\mu+1)}/b_{\nu-1}^{(\mu+1)}, \cdots, b_2^{(\mu+\nu-1)}/b_1^{(\mu+\nu-1)})$$

and

$$M_n^{(\mu)}\leqq b_1^{(n)}+\max(\epsilon, b_{n+1}^{(0)}/b_n^{(0)}, b_n^{(1)}/b_{n-1}^{(1)}, \cdots, b_2^{(n-1)}/b_1^{(n-1)})$$
$$(n=\mu, \mu+1, \cdots).$$

From (2.5) there exists a positive number λ_1, such that $M_nR\leqq\lambda_1<1$ for all $n\geqq n_0$ where n_0 is a suitably chosen number. But $M_n^{(\mu)}\leqq M_n+\epsilon$ so that $M_n^{(\mu)}R\leqq\lambda_1+\epsilon R<\lambda<1$ for $n=\mu+\nu\geqq n_0$ and for ϵ sufficiently small. Now $\sum_{\nu=0}^{\infty}M_{\mu+\nu-1}^{(\mu)}M_{\mu+\nu-2}^{(\mu)}\cdots M_{\mu}^{(\mu)}R^{\nu}$ converges for all μ since the ratio of the

$(\nu+1\text{ts})$ to the νth term is $M_{\mu+\nu}^{(\mu)}R$ which is less than λ for $\nu>n_0$. Moreover, if $\mu \geqq n_0$ this last series is dominated by the series $\sum_{\nu=0}^{\infty}\lambda^\nu=(1-\lambda)^{-1}$. Hence (2.2) converges if the series

$$(1-\lambda)^{-1}\sum_{\mu=n_0}^{\infty}|\gamma_\mu|\,\delta_\mu^{(\mu)}R^\mu$$

converges. This in turn is dominated by

$$(2.7)\qquad (1-\lambda)^{-1}\sum_{\mu=n_0}^{\infty}|\gamma_\mu|\,(1+\epsilon)^\mu R^\mu.$$

But $f(x)$ is analytic at $x=R$ and $\limsup_{\mu\to\infty}|\gamma_\mu|^{1/\mu}R<1$. Therefore (2.7) converges if ϵ is chosen sufficiently small and hence the series (2.2) converges. Theorem II is then proved.

If we specialize Theorem II *by taking* $k_n\equiv 0$, *then* $b_s^{(n)}$ *is a bound for the single coefficient* $a_s^{(n)}$. *The resulting theorem we shall denote by Theorem* III.*

Theorem III is a generalization of the theorem of Okada referred to in the introduction. In the latter theorem condition (2.5) is replaced by

$$0<R<\liminf_{n\to\infty}\left[b_1^{(n)}+\max\left(1/\rho,\ b_{n+1}^{(0)}/b_n^{(0)},\ b_n^{(1)}/b_{n-1}^{(1)},\ \cdots,\ b_2^{(n-1)}/b_1^{(n-1)}\right)\right]^{-1}$$

and condition (2.6) by $1+|a_1^{(n)}|R+|a_2^{(n)}|R^2+\cdots\leqq Q$.

We now point out instances where Theorem II yields results which cannot be established by Theorem III. Consider the set of functions $F_n(x)=x^n(1+a^{(n)}x)(1-x^2)^{-1}$ where $a_{2s}^{(n)}=1$ and $a_{2s-1}^{(n)}=a^{(n)}$ for all s, and $\lim_{n\to\infty}a^{(n)}=1$. Then for a function $f(x)$ analytic for $|x|<1$, using $k_n=-1$ for all n, Theorem II establishes an expansion for $|x|<1$. However, Theorem III would yield at most $|x|<\frac{1}{2}$.

Further Theorem II may be applied with ease to the class of sets of functions which have every other coefficient zero. Application of Theorem II to obtain Neumann's† expansion in Bessel coefficients is immediate if we choose $k_n=(n)^{-1/2}$, or if we choose $k_n=\epsilon$, where ϵ is arbitrarily small.

Theorem III possesses more than a formal advantage over the theorem of Okada. In case $\limsup_{n\to\infty}(M_n-b_1^{(n)})=\infty$, neither theorem yields any results. Hence we assume $M_n-b_1^{(n)}$ bounded for all n. Now consider any

* In (2.5) of Theorem II and Theorem III any or all of the outstanding b's or of ratios of b's with a superscript μ_1 such that $a_\nu^{(\mu_1)}+k_{\mu_1+\nu}a_{\nu-1}^{(\mu_1)}=0$ for all ν may be omitted. If $a_\nu^{(\mu)}+k_{\mu+\nu}a_{\nu-1}^{(\mu)}=0$ for all ν and μ then the only restriction on R is $R<\rho\leqq r$.

† *Theorie der Bessel'schen Functionen*, Leipzig, 1867, pp. 33–35. Also compare with Okada, loc. cit., p. 332.

set of functions $F_n(x)$ such that $\lim\sup_{n\to\infty}|a_1^{(n)}|\neq 0$ and an R such that $1+|a_1^{(n)}|R+|a_2^{(n)}|R^2+\cdots\leqq Q$. Then there exists a positive number $\rho<R$ and also less than $1/(M_n-b_1^{(n)})$ for all n, and it is easily shown that Theorem III will establish a larger region of convergence than the Okada theorem for the expansion of any analytic function $f(x)$ whose nearest singularity is on the circle $|x|=\rho$. In particular, the expansion of $f(x)$, analytic for $\rho<\frac{1}{2}$, in the set of functions $F_n(x)=x^n(1+a^{(n)}x)(1-x^2)^{-1}$ as defined above is established for $|x|<\frac{1}{2}$ by Theorem III and for $|x|<\frac{1}{3}$ by Okada's theorem.

III. Forms of Theorem III using the Cauchy bounds

We shall state here two sets of parallel theorems since in the development of a theory for expansions in products of analytic functions in §VI, it is convenient to have the hypotheses stated in terms of a bound on the functions $|F_n(x)/x^n|$ rather than on the coefficients $a_s^{(n)}$. It is also convenient in order to emphasize this to use the notation $g_n(x)=x^n(1+h_n(x))$ instead of $F_n(x)$. The functions $h_n(x)$ shall be assumed to satisfy the following conditions which we denote by

Conditions (3.1). *$h_n(x)$ vanishes at $x=0$, and $h_n(x)$ is analytic for $|x|<r$.*

Corresponding to any set of functions $h_n(x)$ and to a given positive N less than r, there exists an $M_{N,n}$ such that $|h_n(x)|\leqq M_{N,n}$ for $|x|\leqq N$.

We shall hereafter indicate by K_N $[\overline{K}_N]$ the expression $\lim\sup_{n\to\infty}M_{N,n}$ $[\lim\sup_{n\to\infty}\overline{M}_{N,n}]$ and by N a positive number less than r.

The set of functions $g_n(x)$ satisfying conditions (3.1) belongs to one of two classes:

to A_r if K_N is finite for every N;

to B_r if K_N is infinite for some N.

Moreover, the set of functions $F_n(x)$ shall be said to belong to the class $\overline{A}_r$ if $|a_s^{(n)}|\leqq\overline{M}_{N,n}N^{-s}$ and $\lim\sup_{n\to\infty}\overline{M}_{N,n}\equiv\overline{K}_N$ is finite for every N.

Theorem IV.* *If the set of functions $g_n(x)$ belongs to the class A_r, then any function $f(x)$ analytic for $|x|<\rho$ has a unique uniformly convergent expansion*

$$(3.2)\quad f(x)=\sum_0^\infty \alpha_n g_n(x) \text{ for } |x|\leqq R<G\equiv\min\left(\rho,\ \max_N N(1+K_N)^{-1}\right)$$

and the expansion converges absolutely for $|x|<G$.

Theorem IV′. *If the functions $F_n(x)$ belong to the class $\overline{A}_r$, then any function $f(x)$ analytic for $|x|<\rho$ has a unique uniformly convergent expansion of the form (1) for*

* Theorem IV is essentially equivalent to a theorem due to Takenaka (loc. cit.). He states that K_N is positive but the case $K_N=0$ need not be excluded.

$$|x| \leqq R < \overline{G} \equiv \min\left[\rho, \max_N N(1+\overline{K}_N)^{-1}\right]$$

and absolutely convergent for $|x| < \overline{G}$.*

The proof of Theorem IV′ follows at once from Theorem III. It is evident that $F_n(x)$ is analytic for $|x| < r$ since

$$\limsup_{s\to\infty}[\overline{M}_{N,n}N^{-s}]^{1/s} = N^{-1}.$$

The evaluation of (2.5) using $k_n = 0$ gives $0 < R < N(1+\overline{K}_N)^{-1}$. For such a range of values of R (2.6) would be satisfied since $\overline{M}_{N,n}$ is bounded with respect to n. Now Theorem IV follows at once from Theorem IV′. For if we write $h_n(x) = a_1^{(n)}x + a_2^{(n)}x^2 + a_3^{(n)}x^3 + \cdots$, then from the Cauchy inequalities $|a_s^{(n)}| \leqq M_{N,n}N^{-s}$.

It is clear that if in either case $M_{N,n} \leqq P_n$, then the expansion converges for $|x| < \min[\rho, r(1+\limsup_{n\to\infty}P_n)^{-1}]$.

If in Theorems IV and IV′, $K_N = 0$ for every $N < \rho \leqq r$ then the expansion of $f(x)$ is valid for $|x| < \rho$. The importance of this property justifies the explicit statement of the theorems for this subclass of sets of functions. First, however, we shall state without proof a lemma which will be of value here and also in our later development of expansions in products.

LEMMA 1. *The totality of sets of functions of the class* A_ρ *where* $K_N = 0$ *for every* $N < \rho \leqq r$ *is identical with the totality of sets of functions of the class* $\overline{A}_\rho$ *where* $\overline{K}_N = 0$ *for every* $N < \rho$.

THEOREM V† [V′]. *If the set of functions* $g_n(x)$ $[F_n(x)]$ *belongs to the class* A_ρ $[\overline{A}_\rho]$ *and* $K_N = 0$ $[\overline{K}_N = 0]$ *for every* $N < \rho \leqq r$ *then any function* $f(x)$ *analytic for* $|x| < \rho$ *has an absolutely and uniformly convergent expansion of the form* (3.2) [(1)] *valid for any region* $|x| \leqq R < \rho$. *Further, there is only one such expansion converging uniformly to* $f(x)$ *in a region* $|x| \leqq R'$, $R' > 0$.

From Lemma I, Theorems V and V′ are equivalent.

To obtain the expansion theorem of Valiron‡ from Theorem V′, put $F_n(x) = n!\phi_n(x)$, where $\phi_n(x)$ is an integral function of the form

* Graesser's expansion theorem (loc. cit.) is obtained from Theorem IV′ by replacing K_N by M_N where M_N is the upper bound of $M_{N,n}$ with respect to n. Every case coming under Theorem IV′ can also be treated under Graesser's theorem but IV′ will in general give a larger region of validity for the expansion. The expansion theorems of Pincherle and Graesser are readily shown to be equivalent. See also the work of Izumi and Narumi referred to previously.

† This theorem was proved by Takahashi (1932) and the special case where $M_{N,n}$ goes to zero as $1/n$ by Widder (1929).

‡ Loc. cit. This particular expansion was previously established by Okada (loc. cit., p. 331) for the more general case where $\phi_n(x)$ is analytic for $|x| < r$. See also E. T. Whittaker, *A Course of Modern Analysis*, 1902, p. 110.

$$\phi_n(x) = \frac{x^n}{n!} + a_1\frac{x^{n+1}}{(n+1)!} + \cdots + a_s\frac{x^{n+s}}{(n+s)!} + \cdots .$$

Then

$$|a_s^{(n)}| = \left|\frac{a_s}{(n+1)(n+2)\cdots(n+s)}\right| \leqq A_N N^{-s}(n+1)^{-1},$$

where N is arbitrary.

As an application of Theorem V let

$$g_0(x) = 1, \qquad g_n(x) = x^n f_n[\phi^{p_n}(x)] \qquad (n = 1, 2, \cdots),$$

where $f_n(u)$ is analytic for $|u| < R_1^{p_n}$, $\phi(x)$ is analytic for $|x| < R_2$, $\phi(x)$ is not a constant, $\phi(0) = 0$, $f_n(0) = 1$, and the sequence of positive integral numbers $p_n \to \infty$. Using the notation of Theorem V, write $g_n(x) = x^n(1+G_n(u)) = x^n(1+h_n(x))$, where $u = \phi^{p_n}(x)$. Let $M_n(\sigma^{p_n})$ denote the maximum of $|G_n(u)|$ on the circle $|u| = \sigma^{p_n}$, σ' the upper bound of values of σ such that $\sigma \leqq R_1$ and $M_n(\sigma^{p_n})$ is bounded, δ' the upper bound of values of δ such that $\delta \leqq R_2$ and maximum of $|\phi(x)| \leqq \sigma'$ for $|x| = \delta$. It is easy to prove that* $\lim_{n\to\infty} M_n(\sigma^{p_n}) = 0$ for $\sigma < \sigma'$. Hence in any region $|x| \leqq \delta < \delta'$, the $\lim_{n\to\infty} |h_n(x)| = 0$. Further $g_n(x)$ is analytic in the region $|x| \leqq \delta$. Hence any function $f(x)$ analytic for $|x| < \rho$ has an absolutely and uniformly convergent expansion of the form

$$f(x) = \sum_{n=0}^{\infty} c_n x^n f_n[\phi^{p_n}(x)] \text{ for } |x| \leqq R < \min(\delta', \rho).$$

The special case where $f_n(u)$ and $\phi(x)$ are integral was obtained by Onofri.† In some cases he obtains a region of validity greater than this but the expression which determines this region involves the coefficients c_n.

The above results obtained by Onofri are a generalization of the work of Julia. Julia‡ takes $\phi(x) = x$, $p_n = n$, $f_0(u) = 1$, and $f_n(u) = f_1(u)$ $(n = 1, 2, \cdots)$.

THEOREM VI.§ *Let $a_k^{(n)}(x)$ be a finite or infinite set of functions each analytic for $|x| < \rho$ such that $|a_k^{(n)}(x)| \leqq M_{N,k}$ for $|x| \leqq N$ and every $N < \rho$. Let λ_k be a non-decreasing sequence of positive constants with the limit infinity. If*

$$M_{N,1}\frac{N}{\lambda_1} + M_{N,2}\frac{N^2}{\lambda_1\lambda_2} + M_{N,3}\frac{N^3}{\lambda_1\lambda_2\lambda_3} + \cdots$$

---

* Onofri, loc. cit., p. 211.

† Loc. cit.

‡ Loc. cit.

§ Theorem VI reduces to a theorem due to Carmichael (loc. cit.), when $a_k^{(n)}(x) = a_k(x)$. While Theorem VI is proved as a special case of Theorem V, conversely, Theorem V is a special case of Theorem VI obtained by putting $a_1^{(n)}(x)x/\lambda_{n+1} = h_n(x)$ and $a_r^{(n)}(x) = 0$ for $r = 2, 3, \cdots$.

converges to a sum M_N for every $N<\rho$, then any function $f(x)$ analytic for $|x|<\rho$ has an expansion of the form

$$f(x) = \sum_0^\infty c_n x^n \left\{ 1 + a_1^{(n)}(x)\frac{x}{\lambda_{n+1}} + a_2^{(n)}(x)\frac{x^2}{\lambda_{n+1}\lambda_{n+2}} + \cdots \right\}.$$

This expansion converges uniformly and absolutely in any region $|x| \leqq R<\rho$ and is valid for $|x|<\rho$. Further, there is only one such expansion converging uniformly to $f(x)$ in a region $|x| \leqq R'$, $R'>0$.

Now

$$g_n(x) = x^n\left(1 + a_1^{(n)}(x)\frac{x}{\lambda_{n+1}} + a_2^{(n)}(x)\frac{x^2}{\lambda_{n+1}\lambda_{n+2}} + \cdots\right)$$

and under the hypotheses given $|h_n(x)| \leqq \lambda_1 M_N/\lambda_{n+1}$ which $\to 0$ as $n\to\infty$. Hence the results follow from Theorem V.

Theorems III and IV establish an expansion theory for any set of functions belonging to a class A_ρ for some $\rho\neq 0$. If the set of functions $g_n(x)$ belongs to the class B_ρ for every $\rho\neq 0$, then neither theorem yields any results. Further if $K_N=0$ for every $N<\rho$, then the region of convergence asserted by each theorem is the same, namely, $|x|<\rho$. In case $K_N\neq 0$ the question as to whether by selecting $b_\nu^{(\mu)}$ different from the Cauchy bounds it were possible to establish the validity of the expansion for a larger region by application of Theorem III is left unanswered.

IV. Expansion of an arbitrary analytic function in a class of integral functions

THEOREM VII. *Let λ_n be a sequence of positive numbers such that $\lim_{n\to\infty}\lambda_n = \infty$, and let M be a positive number independent of n and s. If the functions $F_n(x)$ are such that*

$$|a_s^{(n)}| < \frac{M^s}{\lambda_{n+1}\lambda_{n+2}\cdots\lambda_{n+s}} \qquad (n = 0, 1, 2, \cdots; s = 1, 2, 3, \cdots),$$

then any function $f(x)$ analytic for $|x|<\rho$ has a unique expansion of the form (1) valid for $|x|<\rho$ and this expansion converges absolutely and uniformly for values of x such that $|x| \leqq R<\rho$.

We shall make use of Theorem V′ to establish Theorem VII.*

* Theorem VII may be established more easily by direct application of Theorem III. However, we desire to make use in a later discussion of the fact that it follows from Theorem V′ and hence from Theorem V.

The functions $F_n(x)$ are integral functions so it is sufficient to show that corresponding to any N, there exists an M'_N such that

$$\frac{M^s}{\lambda_{n+1}\lambda_{n+2}\cdots\lambda_{n+s}} < \frac{M'_N}{N^s\lambda_{n+1}} \qquad (n = 0, 1, 2, \cdots; s = 1, 2, 3, \cdots).$$

Now $MN/\lambda_{n+s}\to 0$ as $n+s\to\infty$, therefore there exists a q such that $MN/\lambda_{n+s}<1$ for $n+s>q$. But there are only a finite number of values of λ_{n+s} for which $n+s\leqq q$ and hence if we pick M'_N larger than 1 and so large that $(MN)^s/(\lambda_{n+2}\lambda_{n+3}\cdots\lambda_{n+s})<M'_N$, for $n+s\leqq q$, this M'_N will satisfy the desired condition and Theorem VII is established.

To obtain the Neumann* expansion of $f(x)$ in Bessel coefficients we write $F_n(x)=2^n n!J_n(x)$ and obtain the expansion in the set of functions $F_n(x)$ and hence in $J_n(x)$. For s odd $|a_s^{(n)}|=0$ and for s even

$$\begin{aligned} |a_s^{(n)}| &= [2^s(s/2)!(n+1)(n+2)\cdots(n+s/2)]^{-1} \\ &< [(2n+2)(2n+4)\cdots(2n+s)]^{-1} \\ &\leqq (n+1)^{-1/2}(n+2)^{-1/2}\cdots(n+s)^{-1/2}. \end{aligned}$$

In this case $M=1$, and $\lambda_{n+s}=(n+s)^{1/2}$.

We may extend Theorem VII to obtain expansions of the form

$$x^\sigma f(x) = \sum_{n=0}^{\infty} \alpha_n F_{\sigma+n}(x).$$

THEOREM VIII. *Let λ_n be a sequence of positive numbers such that $\lim_{n\to\infty}\lambda_n=\infty$ and let M_s be a positive number independent of n and s. Also let*

$$F_{\sigma+n}(x) = x^{\sigma+n} + \sum_{s=1}^{\infty} a_s^{(n)} x^{\sigma+n+s},$$

where $a_s^{(n)}$ will in general depend on σ, and $|a_s^{(n)}| \leqq M_\sigma^s/(\lambda_{n+1}\lambda_{n+2}\cdots\lambda_{n+s})$. If $f(x)$ is any function analytic for $|x|<\rho$, then $x^\sigma f(x)$ has an absolutely and uniformly convergent expansion of the form

$$x^\sigma f(x) = \sum_{n=0}^{\infty} \alpha_n F_{\sigma+n}(x)$$

valid for $|x|\leqq R<\rho$.

It is sufficient to observe that the function $x^{-\sigma}F_{\sigma+n}(x)$ is of the form $F_n(x)$ as defined in Theorem VII.

---

* Loc. cit., pp. 33–35.

We now apply this theorem to obtain the Gegenbauer† expansion in Bessel functions. Write

$$F_{\sigma+n}(x) = 2^{\sigma+n}\Gamma(\sigma+n+1)J_{\sigma+n}(x) = \sum_{n=0}^{\infty}(-1)^r \frac{\Gamma(\sigma+n+1)x^{\sigma+n+2r}}{2^{2r}r!\Gamma(\sigma+n+r+1)}.$$

Then for s odd $|a_s^{(n)}| = 0$, and for s even

$$|a_s^{(n)}| = [2^s(s/2)!|\sigma+n+1||\sigma+n+2|\cdots|\sigma+n+s/2|]^{-1}.$$

Hence if σ is any point of the complex plane not a negative integer, $|\sigma+n+s|(n+s)^{-1} > Q_\sigma > 0$ and therefore

$$|a_s^{(n)}| < M_\sigma^s(n+1)^{-1/2}(n+2)^{-1/2}\cdots(n+s)^{-1/2}$$

and the Gegenbauer expansion is established.

V. Generalization of expansions in confluent hypergeometric functions

In the preceding section, Theorem VII exhibited a class of sets of functions defined by a condition

$$|a_s^{(n)}| < \phi(n, s) \tag{5.1}$$

such that any set $F_n(x)$ has the following property (which we call Property A): *The expansion of any function $f(x)$ analytic at zero in the functions $F_n(x)$ converges for $|x| < \rho$ where the singularity of $f(x)$ nearest the origin is on the circle $|x| = \rho$.*

We shall now obtain some necessary conditions on $\phi(n, s)$ in order for (5.1) to define a non-null class of sets of functions each set of which possesses Property A. In the first place $\phi(n, s)$ must be positive for all n and s else no functions $F_n(x)$ are determined. Also for every s,

$$\Phi(s) \equiv \lim_{n\to\infty} \inf \phi(n, s) = 0. \tag{5.2}$$

For if not, let ν be the first value of s for which $\Phi(s) > 0$ and let B be a positive number less than 1 and less than $\phi(n, \nu)$ for all n. If $F_p(x) = x^p - B^\nu x^{p+\nu}$, then $\sum_{n=0}^{\infty} B^{n\nu}F_{n\nu}(x)$ converges to 1 if $|x| < 1/B$ and diverges if $|x| > 1/B$. Hence this set $F_n(x)$ does not possess Property A. But $|a_\nu^{(n)}| = B^\nu < \phi(n, \nu)$ and $|a_s^{(n)}| = 0 < \phi(n, s)$ for $s \neq \nu$. Hence no such $\phi(n, s)$ could exist without the condition (5.2).

Now consider any theorem such that for every particular set of functions $F_n(x)$, with coefficients $a_s^{(n)*}$, possessing Property A by virtue of the theorem; there corresponds a positive function $\phi(n, s) \geqq |a_s^{(n)*}|$ such that every set of

† Wiener Sitzungsberichte, (2), vol. 74 (1876), pp. 125–127.

functions $F_n(x)$ for which $|a_s^{(n)}| < \phi(n, s)$ also possesses Property A. We designate such a theorem as of Type O.

Every expansion theorem in which each condition is expressed entirely in terms of an arbitrary set of positive bounds on $a_s^{(n)}$ would be of Type O. Theorem III and all its corollaries are of Type O.†

If there exists an s such that $\liminf_{n\to\infty} a_s^{(n)} \neq 0$, then it is not possible to prove by any theorem of Type O that a particular set of functions $F_n(x)$ possesses Property A. For otherwise there would be determined a class of sets of functions defined by a relation (5.1) such that (5.2) is not satisfied which has just been shown to be impossible. In particular, it is impossible to prove by any such theorem the known fact that Graesser's set of confluent hypergeometric functions‡ possesses Property A.

We next exhibit a general class of sets of functions including this exceptional set of Graesser which possesses Property A.

THEOREM IX. *Let the functions $F_n(x)$ satisfy the conditions*:

$$(5.3) \qquad a_s^{(n)} = \frac{\psi(n+1)\psi(n+2)\cdots\psi(n+s)}{(2cn+d+c)(2cn+d+2c)\cdots(2cn+d+sc)}\,\frac{1}{s!},$$

where c and d are any non-negative real§ constants not both zero and $\psi(n)$ is any function of n such that $|\psi(n)|/(cn+d) < L$, $(n=1, 2, \cdots)$. Then any function $f(x)$ analytic for $|x| < \rho$ has an absolutely and uniformly convergent expansion of the form (1) *valid for $|x| \leqq R < \rho$.*

In Theorem I, take

$$k_{\mu+\nu}^{(\mu)} = \frac{(-1)^\nu \psi(\mu+1)\psi(\mu+2)\cdots\psi(\mu+\nu)}{(2c\mu+d+c\nu)(2c\mu+d+c\nu+c)\cdots(2c\mu+d+2c\nu-c)}\,\frac{1}{\nu!} \qquad (\nu = 1, 2, \cdots).$$

Then‖ $\sum_{t=0}^{\nu} k_{\mu+\nu}^{(\mu+t)} a_t^{(\mu)} = 0$. Let $\delta_{\mu+\nu}^{(\mu)} = |k_{\mu+\nu}^{(\mu)}|$ for $\nu = 0, 1, 2, \cdots$, so that

† The first condition of Theorem III is already expressed by means of any positive bounds $b_s^{(n)}$. This first condition being satisfied for every R, the theorem will yield Property A if and only if $|P_n(x)| < Q_R$ for $|x| = R$ and for every R. But this could happen if and only if $\sum_{s=0}^{\infty} |a_s^{(n)}| R^s < Q_R'$ for every R. In turn this can happen if and only if $\sum_{s=0}^{\infty} [|a_s^{(n)}| + (s!)^{-1}] R^s < Q_R''$ for every R. Given any set of functions with coefficients $a_s^{(n)*}$ and bounds $b_s^{(n)*}$ which yield Property A by this theorem, if we take $\phi(n, s)$ as the smaller of $b_s^{(n)*}$ and $|a_s^{(n)*}| + (s!)^{-1}$ then the conditions for a theorem of Type O are seen to be satisfied.

‡ These functions are obtained from the functions of Theorem IX by taking $\psi(n) = n - k - \frac{1}{2}$, $c=1$, $d=0$, where k is any complex number. In this case $\lim_{n\to\infty} a_1^{(n)} = \frac{1}{2}$.

§ Evidently this restriction may be lightened.

‖ This result follows from the equality

$$\sum_{t=0}^{\nu} (-1)^{\nu-t}(t+1)(t+2)\cdots(t+k)\binom{\nu}{t} = 0 \qquad (k=1, 2, \cdots, \nu-1).$$

$$(5.4)\quad \overline{M}_{\mu+\nu}^{(\mu)} = \frac{\delta_{\mu+\nu+1}^{(\mu)}}{\delta_{\mu+\nu}^{(\mu)}} = \left| \frac{\psi(\mu+\nu+1)(2c\mu+d+c\nu)}{(\nu+1)(2c\mu+d+2c\nu)(2c\mu+d+2c\nu+c)} \right| < \frac{L}{\nu+1},$$

for $\nu = 1, 2, \cdots$, and $\overline{M}_{\mu}^{(\mu)} = |\psi(\mu+1)/(2c\mu+d+c)| < L$. Hence (2.2) is dominated by the series

$$\sum_{\mu,\nu} |\gamma_\mu| \frac{L^\nu R^{\mu+\nu}}{\nu!} = e^{LR} \sum_{\mu=1}^{\infty} |\gamma_\mu| R^\mu.$$

But the last series converges since $f(x)$ is analytic for $|x| \leqq R$. Also for $|x| \leqq R$, $|F_\mu(x)/x^\mu| < e^{LR}$. Hence Theorem IX follows.

Evidently these functions $F_n(x)$ are exceptional in the above sense, that is, $\liminf_{n\to\infty} a_1^{(n)} \neq 0$, provided $|\psi(n)|/(cn+d) > L'$ for an infinite number of values of n.

A simple corollary of Theorem IX would be obtained by taking $c=0$, $d=1$.

VI. Expansions in products of analytic functions

Throughout the present section, we shall make the following assumptions:

A point σ belongs to a region or set of points S of the complex plane.

$g_\sigma(x)$ and $g_\sigma'(x)$ are defined for each point σ of S and are of the form $g_\sigma(x) = x^\sigma(1+h_\sigma(x))$, $g_\sigma'(x) = x^\sigma(1+h_\sigma'(x))$, where h_σ and h_σ' are analytic for $|x| < r$ and vanish at the origin.

N shall denote a positive number less than r.

$f(x)$ is any function analytic for $|x| < \rho$.

κ is any complex number such that $\kappa + n = \sigma_n + \sigma_n'$ for $n = 0, 1, 2, \cdots$, where σ_n and σ_n' are points of S.

$F_\sigma(x)$ and $F_\sigma'(x)$ are defined for each σ of S and are of the form

$$F_\sigma(x) = x^\sigma + \sum_{s=1}^{\infty} a_s^{(\sigma)} x^{\sigma+s}, \qquad F_\sigma'(x) = x^\sigma + \sum_{s=1}^{\infty} \alpha_s^{(\sigma)} x^{\sigma+s}.$$

Theorem X. *If the upper bounds $M_{N,\sigma}$ and $M'_{N,\sigma}$ for $|h_\sigma(x)|$ and $|h_\sigma'(x)|$ respectively in the region $|x| \leqq N$ are such that* $\limsup_{n\to\infty} M_{N,\sigma_n}$ *and* $\limsup_{n\to\infty} M'_{N,\sigma_n}$ *are finite for every N, then $x^\kappa f(x)$ has an absolutely and uniformly convergent expansion of the form*

$$(6.1)\qquad x^\kappa f(x) = \sum_{n=0}^{\infty} c_n g_{\sigma_n}(x) g'_{\sigma_n'}(x) \quad \text{for } |x| \leqq R,$$

where $R < \rho$ and also less than

$$N\left\{1 + \limsup_{n\to\infty} M_{N,\sigma_n} + M'_{N,\sigma_n'} + M_{N,\sigma_n} M'_{N,\sigma_n'}\right\}^{-1}.$$

If we write $x^{-\kappa}g_{\sigma_n}g'_{\sigma'_n}=x^n(1+H_{\kappa+n}(x))$, then the functions $H_{\kappa+n}(x)$ satisfy the conditions in Theorem IV.

THEOREM XI. *If M_{N,σ_n} and M'_{N,σ'_n} defined as in Theorem* X *satisfy the further condition that* $\limsup_{n\to\infty}M_{N,\sigma_n}$ *and* $\limsup_{n\to\infty}M'_{N,\sigma'}$ *are zero for every $N<r=\rho$, then the expansion* (6.1) *is valid for $|x|<\rho$ and the series converges absolutely and uniformly for $|x|\leqq R<\rho$.*

In view of Theorem XI and Lemma 1 we have

THEOREM XII. *Let $F_\sigma(x)$ and $F'_\sigma(x)$ satisfy the conditions*

$$|a_s^{(\sigma)}|\leqq M_{N,\sigma}N^{-s},\qquad |\alpha_s^{(\sigma)}|\leqq M'_{N,\sigma}N^{-s},\qquad (s=1,2,3,\cdots),$$

where $M_{N,\sigma}$ and $M'_{N,\sigma}$ are positive constants. If $\limsup_{n\to\infty}M_{N,\sigma_n}=0$ and $\limsup_{n\to\infty}M'_{N,\sigma_n}=0$ *for every $N<r=\rho$, then $x^\kappa f(x)$ has an absolutely and uniformly convergent expansion of the form*

$$x^\kappa f(x)=\sum_{n=0}^{\infty}c_nF_{\sigma_n}(x)F'_{\sigma'_n}(x)\quad \textit{for}\quad |x|\leqq R<\rho. \tag{6.2}$$

THEOREM XIII. *Let σ be restricted to an unbounded region or set of points S of the complex plane such that $|s(\sigma+s)^{-1}|$ $(s=1,2,\cdots)$ is bounded with respect to both s and σ for σ in S. If $\lim_{n\to\infty}|\sigma_n|=\infty$, $\lim_{n\to\infty}|\sigma_n'|=\infty$, and $F_\sigma(x)$, $F'_\sigma(x)$ are such that $|a_s^{(\sigma)}|$ and $|\alpha_s^{(\sigma)}|$ are bounded by $M^s/(|\sigma+1|^m|\sigma+2|^m\cdots|\sigma+s|^m)$ for all σ in S, M being a positive constant independent of σ and s, and m a fixed positive number; then $x^\kappa f(x)$ has a uniformly and absolutely convergent expansion of the form* (6.2) *for $|x|\leqq R<\rho$.*

We shall make the proof of this theorem depend on Theorem XII. It is sufficient to show under the named conditions that we can determine an $\overline{M}$ corresponding to any N arbitrarily large such that

$$\frac{M^s}{|\sigma_n+1|^m|\sigma_n+2|^m\cdots|\sigma_n+s|^m}\leqq\frac{\overline{M}N^{-s}}{|\sigma_n+1|^m}.$$

Now for σ in S, $s/|\sigma+s|<K$ and hence

$$\frac{(MN)^s}{|\sigma_n+2|^m|\sigma_n+3|^m\cdots|\sigma_n+s^m}<\frac{(MN)^s}{K^{m(s-1)}2^m\cdot3^m\cdots s^m}.$$

But $MNK^{-m}/s^m\to0$ as $s\to\infty$. Therefore

$$\frac{(MN)^s}{|\sigma_n+2|^m|\sigma_n+3|^m\cdots|\sigma_n+s|^m}\leqq\overline{M}.$$

A region S suitable for application of Theorem XIII would be obtained

by excluding the negative integers $-1, -2, \cdots, -j$ by circles of radius l and by excluding any sector containing the negative axis of reals in its interior and radiating from the point $-(j+\frac{1}{2})$. Here j is a positive integer and l is any positive number.

To apply Theorem XIII to the Bessel functions, write

$$F_\sigma(x) = F_\sigma'(x) = \Gamma(\sigma+1)2^\sigma J_\sigma(x)$$
$$= x^\sigma + \sum_{r=1}^{\infty} (-1)^r \frac{x^{\sigma+2r}}{r!2^{2r}(\sigma+1)(\sigma+2)\cdots(\sigma+r)}.$$

Then for s odd $|a_s^{(\sigma)}| = 0$, and for s even

$$|a_s^{(\sigma)}| = \left((s/2)!2^s|\sigma+1||\sigma+2|\cdots|\sigma+s/2|\right)^{-1}$$
$$\leqq M^s(|\sigma+1||\sigma+2|\cdots|\sigma+s|)^{-1/2},$$

where M is a properly chosen positive number and σ is restricted to any region S for which $s/|\sigma+s|$ is uniformly bounded.

The Neumann-Gegenbauer* expansion in products of Bessel functions,

$$x^{\mu+\nu}f(x) = \sum_{n=0}^{\infty} c_n J_{\mu+n/2}(x)J_{\nu+n/2}(x),$$

is obtained from Theorem XIII if we put $F_\sigma(x) = F_\sigma'(x) = \Gamma(\sigma+1)2^\sigma J_\sigma(x)$, $\mu+\nu=\kappa$, $\sigma_n=\mu+n/2$, $\sigma_n'=\nu+n/2$, and restrict μ and ν so that $\mu+n/2$ and $\nu+n/2$ are not negative integers. Then σ_n, σ_n' are in S and $|\sigma_n|$, $|\sigma_n'| \to \infty$ as $n \to \infty$.

If we take $\kappa=0$, $\sigma_n=n/2$, $\sigma_n'=n/2$ for n even and $\sigma_n=(n-1)/2$, $\sigma_n'=(n+1)/2$ for n odd, we obtain the expansion of any function $f(x)$ analytic about zero in the form

$$f(x) = c_0J_0^2(x) + c_1J_0(x)J_1(x) + c_2J_1^2(x) + c_3J_1(x)J_2(x) + \cdots.$$

From this expansion we have at once the Neumann† expansion of an even analytic function and for an odd analytic function: namely, $f(x) = \sum_{n=0}^{\infty} c_n J_n^2(x)$ and $f(x) = \sum_{n=0}^{\infty} c_n J_n(x)J_{n+1}(x)$.

The expansion due to Nielsen‡ of the form

$$x^\nu f(x) = \sum_{n=0}^{\infty} a_n x^{(\nu+n)/2} J_{(\nu+n)/2}(x)$$

---

* See Watson, *Theory of Bessel Functions*, p. 525.

† Berichte der K. Sächsische Gesellschaft der Wissenschaften, vol. 21 (1869), pp. 221–256.

‡ Nyt Tidsskrift for Matematik, vol. 9 (1898), pp. 77–79.

is an immediate consequence of Theorem XIII. Evidently much more general expansions of the form

$$x^{\kappa}f(x) = \sum_{n=0}^{\infty} \alpha_n J_{\sigma_n}(x) x^{\sigma_n'}$$

also follow from Theorem XIII.

Obviously the theorems of this section may be extended immediately to expansions in series each term of which is made up of the product of k functions. We shall state the theorem corresponding to Theorem XIII for expansions in products of k Bessel functions.

THEOREM XIV. *Let σ be restricted to an unbounded region S of the complex plane such that $s/|\sigma+s|$ $(s=1, 2, \cdots)$ is uniformly bounded; and let κ be a complex number such that $\kappa+n=\sigma_n^{(1)}+\sigma_n^{(2)}+\cdots+\sigma_n^{(k)}$, where $\sigma_n^{(1)}, \sigma_n^{(2)}, \cdots, \sigma_n^{(k)}$ are in S and $\lim_{n\to\infty}|\sigma_n^{(r)}| = \infty$ $(r=1, 2, \cdots, k)$. Then if $f(x)$ is any function analytic for $|x|<\rho$, $x^{\kappa}f(x)$ has the absolutely and uniformly convergent expansion*

$$x^{\kappa}f(x) = \sum_{n=0}^{\infty} c_n J_{\sigma_n^{(1)}}(x) J_{\sigma_n^{(2)}}(x) \cdots J_{\sigma_n^{(k)}}(x) \text{ for } |x| \leqq R < \rho.$$

A special case of this theorem which is a generalization of the Neumann-Gegenbauer expansion is

$$x^{\mu}f(x) = \sum_{n=0}^{\infty} c_n J_{\mu_1+\theta_1 n}(x) J_{\mu_2+\theta_2 n}(x) \cdots J_{\mu_k+\theta_k n}(x),$$

where μ_r and θ_r $(r=1, 2, \cdots, k)$ are any complex numbers subject to the conditions that $\mu_r+\theta_r n$ is not a negative integer, θ_r does not vanish, $\mu=\mu_1+\mu_2+\cdots+\mu_k$, and $\theta_1+\theta_2+\cdots+\theta_k=1$. Here $f(x)$ is analytic for $|x|<\rho$ and the expansion converges absolutely and uniformly for $|x|\leqq R<\rho$.

It is not necessary that each term of an expansion be the product of exactly k functions. Under appropriate hypotheses the theorems of this section may be generalized in such a way that we obtain expansions where any term may be the product of one, two, up to k functions so long as k is bounded for all terms.

UNIVERSITY OF ILLINOIS,
URBANA, ILL.

BERNSTEIN'S THEOREM AND TRIGONOMETRIC APPROXIMATION*

BY

DUNHAM JACKSON

1. **Introduction.** In a recent paper† the writer has presented a sequence of theorems relating to normalized polynomials and polynomial approximation together with corresponding results for normalized trigonometric sums and trigonometric approximation over an entire period. The present paper is concerned with similar questions in connection with trigonometric sums considered over a part of a period, as well as some additional observations on the polynomial case. The treatment is based largely on Bernstein's theorem on the derivative of a trigonometric sum, and incidentally involves or suggests certain modifications and adaptations of that theorem, both for trigonometric sums and for polynomials.

2. **Standard theorems.** The fundamental theorem of Bernstein for trigonometric sums may be stated as follows.‡

THEOREM B1. *If $T_n(x)$ is a trigonometric sum of the nth order such that $|T_n(x)| \leqq L$ for all real values of x, then $|T_n'(x)| \leqq nL$ for all real values of x.*

From this can be deduced immediately§ the corresponding theorem for polynomials:‖

THEOREM B2. *If $P_n(x)$ is a polynomial of the nth degree such that $|P_n(x)| \leqq L$ for $-1 \leqq x \leqq 1$, then $|P_n'(x)| \leqq nL/(1-x^2)^{1/2}$ for $-1 < x < 1$.*

By a linear transformation of the independent variable the theorem can be generalized to an arbitrary interval (a, b), to read thus:

---

* Presented to the Society, under two different titles, September 11, 13, 1935; received by the editors, October 15, 1935.

† *Certain problems of closest approximation*, Bulletin of the American Mathematical Society, vol. 39 (1933), pp. 889–906.

‡ S. Bernstein, *Sur l' ordre de la meilleure approximation des fonctions continues par des polynomes de degré donné*, Mémoire couronné, Brussels, 1912, p. 20 (with $2nL$ instead of nL as upper bound for $|T_n'(x)|$); M. Riesz, *Eine trigonometrische Interpolationsformel und einige Ungleichungen für Polynome*, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 23 (1914), pp. 354–368; pp. 357, 360–361; C. de la Vallée Poussin, *Sur le maximum du module de la dérivée d'une expression trigonométrique d'ordre et de module bornés*, Comptes Rendus, vol. 166 (1918), pp. 843–846; also, e.g., the writer's Colloquium, *The Theory of Approximation*, New York, 1930, pp. 80–81.

§ See, e.g., Colloquium, p. 92.

‖ S. Bernstein, op. cit., pp. 6–11.

THEOREM B2a. *If $P_n(x)$ is a polynomial of the nth degree such that $|P_n(x)| \leqq L$ for $a \leqq x \leqq b$, then*

$$|P_n'(x)| \leqq \frac{nL}{[(b-x)(x-a)]^{1/2}}$$

for $a < x < b$.

A closely related theorem for polynomials is that of Markoff:*

THEOREM M1. *If $P_n(x)$ is a polynomial of the nth degree such that $|P_n(x)| \leqq L$ for $-1 \leqq x \leqq 1$, then $|P_n'(x)| \leqq n^2 L$ throughout the same interval.*

The generalized statement for an arbitrary interval is:

THEOREM M1a. *If $P_n(x)$ is a polynomial of the nth degree such that $|P_n(x)| \leqq L$ for $a \leqq x \leqq b$, then $|P_n'(x)| \leqq 2n^2 L/(b-a)$ throughout the interval.*

In a somewhat weaker form,† which is nevertheless sufficient for present purposes, namely, with $2n^2L$ as upper bound for $|P_n'(x)|$ instead of n^2L, Theorem M1 can be obtained as an easy corollary of Theorem B1, and Theorem M1a follows with coefficient 4 in place of 2.

Less frequently cited is the following theorem, which is due also to Bernstein:‡

THEOREM B3. *If $P_{n-1}(x)$ is a polynomial of degree $n-1$ such that*

$$|(1-x^2)^{1/2} P_{n-1}(x)| \leqq L$$

for $-1 \leqq x \leqq 1$, then

$$\left| \frac{d}{dx}[(1-x^2)^{1/2} P_{n-1}(x)] \right| \leqq \frac{nL}{(1-x^2)^{1/2}}$$

for $-1 < x < 1$.

For a general interval this becomes

THEOREM B3a. *If $P_{n-1}(x)$ is a polynomial of degree $n-1$ such that*

$$|(b-x)^{1/2}(x-a)^{1/2} P_{n-1}(x)| \leqq L$$

for $a \leqq x \leqq b$, then

* See, e.g., S. Bernstein, op. cit., pp. 11–13; M. Riesz, loc. cit., pp. 359–360.

† See D. Jackson, *On the convergence of certain trigonometric and polynomial approximations*, these Transactions, vol. 22 (1921), pp. 158–166; pp. 162–163. (In the displayed formula midway between (4) and (5) on p. 163 the absolute value bars inclosing the expression $Q_n(\theta)/\sin\theta$ should be omitted.)

‡ S. Bernstein, op. cit., pp. 17–19; see also D. Jackson, *A generalized problem in weighted approximation*, these Transactions, vol. 26 (1924), pp. 133–154; pp. 140–141.

$$\left| \frac{d}{dx} \left[(b - x)^{1/2}(x - a)^{1/2} P_{n-1}(x) \right] \right| \leqq \frac{nL}{[(b - x)(x - a)]^{1/2}}$$

for $a < x < b$.

3. **Variations.** If an upper bound is given for the absolute value of a trigonometric sum over a part of a period, a conclusion is obtained resembling that of Theorem B2a:

THEOREM 1. *If* $T_n(x)$ *is a trigonometric sum of the nth order such that* $|T_n(x)| \leqq L$ *for* $a \leqq x \leqq b$, *where** $0 < b - a < 2\pi$, *then*

$$|T_n'(x)| \leqq \frac{CnL}{[(b - x)(x - a)]^{1/2}}$$

for $a < x < b$, *where* C *depends only on* a *and* b.

Throughout this paper C will be used as a general notation for a constant, differing in general from one formula to another, and in any particular instance depending on the range assigned to the independent variable, but not depending on anything else. The conclusion of Theorem M1a, for example, has the form $|P_n'(x)| \leqq Cn^2L$. Where the magnitude of the constant is not determined its precise value is immaterial for the purpose in hand.

Theorem 1 has been proved elsewhere.† An alternative form of demonstration will be presented here, to serve as a pattern for other proofs which are to follow.

Let $C_n(x)$ be a cosine sum of the nth order such that $|C_n(x)| \leqq L$ for $a \leqq x \leqq b$, where $0 < a < b < \pi$. Then $C_n(x)$ is a polynomial of the nth degree in $z = \cos x$, say $p_n(z)$, and $|p_n(z)| \leqq L$ for $B \leqq z \leqq A$, if $A = \cos a$, $B = \cos b$. According to Theorem B2a, therefore,

$$|p_n'(z)| \leqq \frac{nL}{[(A - z)(z - B)]^{1/2}}$$

in the interior of the interval. By the mean value theorem, inasmuch as $\sin x$ has a positive lower bound for $a \leqq x \leqq b$, $A - z \geqq C(x - a)$, and $z - B \geqq C(b - x)$. Consequently

$$|C_n'(x)| = |\sin x \, p_n'(z)| \leqq \frac{CnL}{[(b - x)(x - a)]^{1/2}}$$

in the interval considered.

Let $S_n(x)$ be a sine sum of the nth order satisfying the condition that

* If $b - a \geqq 2\pi$, Theorem B1 is applicable.

† D. Jackson, these Transactions, vol. 26, loc. cit., pp. 141–145.

$|S_n(x)| \leqq L$ throughout the same interval. It is possible to write $S_n(x)$ as the product of $\sin x$ by a polynomial $q_{n-1}(z)$ of degree $n-1$, and since $\sin x$ has a positive lower bound in the interval $|q_{n-1}(z)| \leqq CL$. Hence

$$|q'_{n-1}(z)| \leqq \frac{C(n-1)L}{[(A-z)(z-B)]^{1/2}} \leqq \frac{CnL}{[(b-x)(x-a)]^{1/2}},$$

with different values of C in the second and third members, and

$$\begin{aligned} |S'_n(x)| &= |\cos x\, q_{n-1}(z) - \sin^2 x\, q'_{n-1}(z)| \\ &\leqq CL + \frac{CnL}{[(b-x)(x-a)]^{1/2}} \leqq \frac{CnL}{[(b-x)(x-a)]^{1/2}}, \end{aligned}$$

again with different values of C.

Now let $T_n(x)$ be an arbitrary trigonometric sum of the nth order such that $|T_n(x)| \leqq L$ for $a \leqq x \leqq b$ *and also* for $-b \leqq x \leqq -a$, still with $0 < a < b < \pi$. Then L is an upper bound likewise for the absolute values of

$$C_n(x) = \tfrac{1}{2}[T_n(x) + T_n(-x)], \qquad S_n(x) = \tfrac{1}{2}[T_n(x) - T_n(-x)],$$

in the interval (a, b). So $|C'_n(x)|$ and $|S'_n(x)|$ have upper bounds of the form obtained in the preceding paragraphs, for $a < x < b$. And as $|C'_n(x)|$ and $|S'_n(x)|$ are even functions,

$$|C'_n(x)| = |C'_n(-x)| \leqq \frac{CnL}{[(b+x)(-x-a)]^{1/2}} = \frac{CnL}{[(a'-x)(x-b')]^{1/2}}$$

for $b' < x < a'$, if $a' = -a$, $b' = -b$, and $S'_n(x)$ is similarly bounded. So $|T'_n(x)| = |C'_n(x) + S'_n(x)|$ has upper bounds of the form

$$\frac{CnL}{[(b-x)(x-a)]^{1/2}}, \qquad \frac{CnL}{[(a'-x)(x-b')]^{1/2}}$$

in the intervals (a, b) and (b', a') respectively.

More generally, let $T_n(x)$ be a trigonometric sum of the nth order having L as an upper bound for its absolute value in the intervals $a_1 \leqq x \leqq b_1$ and $a_2 \leqq x \leqq b_2$, where (a_1, b_1) and (a_2, b_2) are any two intervals of equal length, having no point in common, and both contained within a period:

$$b_1 - a_1 = b_2 - a_2, \qquad a_1 < b_1 < a_2 < b_2 < a_1 + 2\pi.$$

The change of variable $s = x - \frac{1}{2}(b_1 + a_2)$ transforms $T_n(x)$ into a trigonometric sum of the nth order in s with absolute value $\leqq L$ throughout two equal intervals symmetrically situated with respect to the origin. To this sum $\tau_n(s)$ the discussion of the preceding paragraph is applicable. Since $\tau'_n(s) = T'_n(x)$,

the conclusion as expressed in terms of the variable x is that $|T_n'(x)|$ has the upper bounds

$$\frac{CnL}{[(b_1 - x)(x - a_1)]^{1/2}}, \qquad \frac{CnL}{[(b_2 - x)(x - a_2)]^{1/2}}$$

respectively in (a_1, b_1) and (a_2, b_2).

Finally, *let* $T_n(x)$ *be a trigonometric sum of the nth order such that* $|T_n(x)| \leqq L$ *for* $a \leqq x \leqq b$, *where* (a, b) *is any interval of length* $< 2\pi$. Let the interval (a, b) be covered by four intervals $(a_1, b_1), \cdots, (a_4, b_4)$ of equal length, overlapping two by two but with no point common to any three, extending from a as left-hand extremity of the first interval to b as right-hand extremity of the fourth; specifically, let c_1, c_2, c_3 be the points dividing (a, b) into fourths, in order from left to right, let h be a positive number $< (b-a)/12$, and let

$$a_1 = a, \qquad b_1 = c_1 + 2h; \qquad a_2 = c_1 - h, \qquad b_2 = c_2 + h;$$
$$a_3 = c_2 - h, \qquad b_3 = c_3 + h; \qquad a_4 = c_3 - 2h, \qquad b_4 = b.$$

The preceding paragraph can be applied to the first and third intervals, and again to the second and fourth intervals. If the results are interpreted successively for the interval $a < x \leqq c_1$ (in which $b_1 - x$ has a positive lower bound), for the interval $c_1 \leqq x \leqq c_2$ (in which $(b_2 - x)(x - a_2)$ has a positive lower bound), for the interval $c_2 \leqq x \leqq c_3$ (in which $(b_3 - x)(x - a_3)$ has a positive lower bound), and for the interval $c_3 \leqq x < b$ (in which $x - a_4$ has a positive lower bound), it appears in summary that

$$|T_n'(x)| \leqq \frac{CnL}{[(b - x)(x - a)]^{1/2}}$$

for $a < x < b$, as the assertion of Theorem 1 requires. Since the essential conditions of the problem are unaffected by a change of origin for the independent variable it is clear without re-examination of the details of the proof that the magnitude of C as a function of a and b depends only on the length of the interval, not on the position of its initial point.

A similar method has been used in another passage* to adapt Markoff's theorem to the case of trigonometric sums, with the simplification that since the end points do not have an exceptional status in Markoff's theorem it is not necessary to use overlapping subintervals in dealing with the various parts of the given interval, a simple subdivision of the latter into fourths being all that is required. The conclusion may be repeated here for reference:

* D. Jackson, *On the application of Markoff's theorem to problems of approximation in the complex domain*, Bulletin of the American Mathematical Society, vol. 37 (1931), pp. 883–890; pp. 887–889.

THEOREM 2. *If $T_n(x)$ is a trigonometric sum of the nth order such that $|T_n(x)| \leqq L$ for $a \leqq x \leqq b$, where $0 < b-a < 2\pi$, then*

$$|T_n'(x)| \leqq Cn^2L$$

throughout the same interval, the magnitude of C depending only on the length of the interval.

A trigonometric analogue of Theorem B3a is this:

THEOREM 3. *If $T_{n-1}(x)$ is a trigonometric sum of order $n-1$ such that*

$$|(b-x)^{1/2}(x-a)^{1/2}T_{n-1}(x)| \leqq L$$

for $a \leqq x \leqq b$, where $0 < b-a < 2\pi$, then

$$\left| \frac{d}{dx}[(b-x)^{1/2}(x-a)^{1/2}T_{n-1}(x)] \right| \leqq \frac{CnL}{[(b-x)(x-a)]^{1/2}}$$

for $a < x < b$.

The proof is obtained by successive stages corresponding to those in the proof of Theorem 1. If $C_{n-1}(x)$ is a cosine sum of order $n-1$ such that

$$|(b-x)^{1/2}(x-a)^{1/2}C_{n-1}(x)| \leqq L$$

for $a \leqq x \leqq b$, where $0 < a < b < \pi$, and if $z = \cos x$, $A = \cos a$, $B = \cos b$, $C_{n-1}(x) = p_{n-1}(z)$, Theorem B3a is applicable to the polynomial $p_{n-1}(z)$ in the interval (B, A), and

$$\left| \frac{d}{dz}[(A-z)^{1/2}(z-B)^{1/2}p_{n-1}(z)] \right| \leqq \frac{nL}{[(A-z)(z-B)]^{1/2}}$$

for $B < z < A$. Hence, as $[(b-x)(x-a)]^{1/2}/[(\cos a - \cos x)(\cos x - \cos b)]^{1/2}$ is bounded and has a bounded derivative in the interval under consideration, it follows that

$$\left| \frac{d}{dx}[(b-x)^{1/2}(x-a)^{1/2}C_{n-1}(x)] \right| \leqq \frac{CnL}{[(b-x)(x-a)]^{1/2}}$$

throughout the same interval. A similar result is obtained if $C_{n-1}(x)$ is replaced by a sine sum $S_{n-1}(x)$, of order $n-1$, the differences between this case and that of the cosine sum being taken care of as at the corresponding stage in the proof of Theorem 1.

Let $T_{n-1}(x)$ be any trigonometric sum of order $n-1$ such that

$$|(b-x)^{1/2}(x-a)^{1/2}T_{n-1}(x)| \leqq L$$

for $a \leqq x \leqq b$, and also

$$| (-a-x)^{1/2}(x+b)^{1/2}T_{n-1}(x) | \leqq L$$

for $-b \leqq x \leqq -a$. The latter condition means that

$$| (b-x)^{1/2}(x-a)^{1/2}T_{n-1}(-x) | \leqq L$$

for $a \leqq x \leqq b$, so that the cosine and sine sums

$$C_{n-1}(x) = \tfrac{1}{2}[T_{n-1}(x) + T_{n-1}(-x)], \quad S_{n-1}(x) = \tfrac{1}{2}[T_{n-1}(x) - T_{n-1}(-x)]$$

satisfy the conditions of the paragraph preceding. Consequently

$$\left| \frac{d}{dx}[(b-x)^{1/2}(x-a)^{1/2}T_{n-1}(x)] \right| \leqq \frac{CnL}{[(b-x)(x-a)]^{1/2}}$$

for $a < x < b$, and at the same time, for $-b < x < -a$,

$$\left| \frac{d}{dx}[(-a-x)^{1/2}(x+b)^{1/2}T_{n-1}(x)] \right| \leqq \frac{CnL}{[(-a-x)(x+b)]^{1/2}}.$$

By a change of origin the two intervals can be replaced by any two intervals of equal length contained within a period, and neither overlapping nor having an end point in common. And the proof of Theorem 3 can be completed in the same manner as that of Theorem 1, by the use of four overlapping subintervals covering the interval to which the theorem refers; if

$$| (b-x)^{1/2}(x-a)^{1/2}T_{n-1}(x) | \leqq L$$

for $a \leqq x \leqq b$, and if the same auxiliary notation is used as in the earlier proof, then certainly

$$| (b_1-x)^{1/2}(x-a_1)^{1/2}T_{n-1}(x) | \leqq L$$

for $a_1 \leqq x \leqq b_1$, and similarly for each of the other subintervals, while the final synthesis of the conclusions pertaining to the various subintervals, though less immediate than in the other case, is complicated only by a multiplicity of simple details, which it would be superfluous to enumerate.

A corollary derivable from Theorems 1 and 3 jointly by the use of overlapping intervals, of which in this case only two are required, is a theorem intermediate between them in form:

THEOREM 4. *If* $T_{n-1}(x)$ *is a trigonometric sum of order* $n-1$ *such that*

$$| (x-a)^{1/2}T_{n-1}(x) | \leqq L$$

for $a \leqq x \leqq b$, *where* $0 < b-a < 2\pi$, *then, for* $a < x < b$,

$$\left| \frac{d}{dx}[(x-a)^{1/2}T_{n-1}(x)] \right| \leqq \frac{CnL}{[(b-x)(x-a)]^{1/2}}.$$

Let $c=\frac{1}{2}(a+b)$, $b_1=c+h$, $a_2=c-h$, with $0<h<\frac{1}{2}(b-a)$. Inasmuch as $(b_1-x)^{1/2}\leqq C$ for $a\leqq x\leqq b_1$, application of Theorem 3 to the interval (a, b_1) gives for $a<x\leqq c$ the relations

$$\left|\frac{d}{dx}\left[(b_1-x)^{1/2}(x-a)^{1/2}T_{n-1}(x)\right]\right| \leqq \frac{CnL}{[(b_1-x)(x-a)]^{1/2}} \leqq \frac{CnL}{(x-a)^{1/2}} \leqq \frac{CnL}{[(b-x)(x-a)]^{1/2}},$$

with different values of C in the successive members of the continued inequality; and hence by further use of the fact that $1/(b_1-x)^{1/2}$ and its derivative are bounded for $a<x\leqq c$ it appears that

$$\left|\frac{d}{dx}\left[(x-a)^{1/2}T_{n-1}(x)\right]\right| \leqq \frac{CnL}{[(b-x)(x-a)]^{1/2}} \tag{1}$$

throughout this interval. In (a_1, b), where $|T_{n-1}(x)|\leqq CL$, Theorem 1 is applicable, and for $c\leqq x<b$ the conclusion can be interpreted to give a relation of the form (1). So the assertion of the theorem is valid for the entire interval (a, b).

Naturally the ends of the interval can be interchanged in the statement, the factor $(x-a)^{1/2}$ being replaced by $(b-x)^{1/2}$.

A corresponding theorem for polynomials can be similarly obtained by combination of Theorems B2a and B3a, or deduced from the trigonometric case by means of a cosine substitution:

THEOREM 5. *If $P_{n-1}(x)$ is a polynomial of degree $n-1$ such that*

$$|(x-a)^{1/2}P_{n-1}(x)|\leqq L$$

for $a\leqq x\leqq b$, then

$$\left|\frac{d}{dx}\left[(x-a)^{1/2}P_{n-1}(x)\right]\right| \leqq \frac{CnL}{[(b-x)(x-a)]^{1/2}}$$

for $a<x<b$.

The same procedure leads to a proposition relating to the behavior of a trigonometric sum over an entire period.

THEOREM 6. *If $T_{n-1}(x)$ is a trigonometric sum of order $n-1$ such that*

$$|\sin^{1/2}(x/2)T_{n-1}(x)|\leqq L$$

for $0\leqq x\leqq 2\pi$, then

$$\left|\frac{d}{dx}\left[\sin^{1/2}(x/2)T_{n-1}(x)\right]\right| \leqq \frac{AnL}{\sin^{1/2}(x/2)}$$

for $0<x<2\pi$, where A is an absolute constant.

It is clear that this is equivalent to a superficially more general statement with $[\sin \frac{1}{2}(x-a)]^{1/2}$ in place of $(\sin \frac{1}{2}x)^{1/2}$, and with the interval from 0 to 2π replaced by that from a to $a+2\pi$. Also it is only for reasons of convenience that $\sin \frac{1}{2}x$ is used rather than an algebraic expression vanishing in the same manner. The same letter A will be used to denote a number of different absolute constants in succession, each in its place determinate once for all.

In consequence of the hypothesis $|x^{1/2}T_{n-1}(x)| \leqq AL$ for $0 \leqq x \leqq 3\pi/2$. (Any other fixed number between π and 2π would serve equally well.) By Theorem 4

$$\left| \frac{d}{dx} [x^{1/2}T_{n-1}(x)] \right| \leqq \frac{AnL}{[x(3\pi/2 - x)]^{1/2}}$$

for $0 < x < 3\pi/2$, the symbol C being replaced by that for an absolute constant because the interval is completely determinate. Here $x^{1/2}$ in the left-hand member can be replaced by $\sin^{1/2}(x/2)$ without changing the form of the upper bound on the right, while

$$\frac{AnL}{[x(3\pi/2 - x)]^{1/2}} \leqq \frac{AnL}{x^{1/2}} \leqq \frac{AnL}{\sin^{1/2}(x/2)}$$

for $0 < x \leqq \pi$. Replacement of x by $2\pi - x$ leads to a similar conclusion for the interval $(\pi, 2\pi)$, and so completes the proof of the theorem.

After successive consideration of $P_n(x)$ and $(1-x^2)^{1/2}P_{n-1}(x)$ it may be pointed out that the function $(1-x^2)P_{n-2}(x)$ (the symbol $P_{n-2}(x)$ being regarded as self-explanatory) is merely a polynomial of the nth degree (specialized, to be sure, by the circumstance that it vanishes at the points $+1$ and -1), and as such comes under Theorem B2, if an upper bound is given for its absolute value in $(-1, 1)$. The fact that the vanishing of the polynomial for $x = \pm 1$ does not affect the order of magnitude of the upper bound obtained for its derivative as x approaches the ends of the interval is illustrated by taking for $(1-x^2)P_{n-2}(x)$ the function $1-\cos n\theta$, with $x = \cos \theta$, for even values of n.

The expression $(b-x)(x-a)T_{n-2}(x)$ of course does not represent a trigonometric sum. It is, however, essentially equivalent, as far as the properties involved in this discussion are concerned, to $\sin \frac{1}{2}(b-x) \sin \frac{1}{2}(x-a)T_{n-2}(x)$, which is a trigonometric sum of order $n-1$.

The hypothesis that $|(1-x^2)^{-1/2}P_n(x)| \leqq L$ in $(-1, 1)$ does lead at once to a different form of conclusion, not with regard to the derivative of the product but with regard to the derivative of $P_n(x)$ itself. Since

$$|P_n(x)| \leqq L(1 - x^2)^{1/2}$$

it follows that $P_n(x)$ vanishes for $x=\pm 1$, and must be of the form $(1-x^2)p_{n-2}(x)$, where $p_{n-2}(x)$ is a polynomial of degree $n-2$. From the relation

$$\left| (1-x^2)^{1/2}p_{n-2}(x) \right| \leqq L$$

it appears by Theorem B3 that

$$\left| \frac{d}{dx}\left[(1-x^2)^{1/2}p_{n-2}(x)\right] \right| \leqq \frac{(n-1)L}{(1-x^2)^{1/2}} \leqq \frac{nL}{(1-x^2)^{1/2}}.$$

Hence

$$\begin{aligned} \left| P_n'(x) \right| &= \left| \frac{d}{dx}\left\{(1-x^2)^{1/2}\left[(1-x^2)^{1/2}p_{n-2}(x)\right]\right\} \right| \\ &\leqq (1-x^2)^{1/2}\left| \frac{d}{dx}\left[(1-x^2)^{1/2}p_{n-2}(x)\right] \right| \\ &\quad + \frac{|x|}{(1-x^2)^{1/2}}\left| (1-x^2)^{1/2}p_{n-2}(x) \right| \\ &\leqq nL + \frac{L}{(1-x^2)^{1/2}} \end{aligned}$$

for $-1<x<1$, where in the last member the factor n and the factor $(1-x^2)^{-1/2}$ do not occur in the same term.

On the assumption that $\left|(1-x^2)^{-1}P_n(x)\right| \leqq L$ for $-1<x<1$ it is possible by writing

$$P_n(x) = (1-x^2)p_{n-2}(x), \qquad \left| p_{n-2}(x) \right| \leqq L,$$

and applying Bernstein's theorem to $p_{n-2}(x)$ to obtain the relation

$$\left| P_n'(x) \right| \leqq nL(1-x^2)^{1/2} + 2L.$$

This type of argument will not be carried further here.

4. **Upper bounds for normalized trigonometric sums and polynomials.** The theorems of this section and the next are closely associated in content and arrangement with those of the paper *Certain problems of closest approximation*, previously cited, and henceforth to be referred to simply by the letter A. They are related to the preceding section, not by direct dependence on the results of that section, with the exception of Theorem 2, but by the use of the method of overlapping intervals in making the transition from the case of polynomials to that of trigonometric sums. The conclusions of the

present section have a bearing on the theory set forth in the writer's paper* on *Orthogonal trigonometric sums*. A system of trigonometric sums normalized and orthogonal over an interval (a, b) of length less than 2π can of course be regarded as constituting a system of the same character for an entire period, with weight function vanishing outside (a, b).

An exponent s will be of frequent occurrence. In analogy with the use of the symbol C above, C_s will be used to denote successively a number of different constants, each depending on the interval specified for the independent variable, and on s, but not depending on anything else.

Let $T_n(x)$ be an arbitrary trigonometric sum of the nth order, let s be an arbitrary positive number, and let

$$H_{ns} = \int_a^b | T_n(x) |^s dx,$$

the length of the interval (a, b) being less† than 2π. An upper bound for the absolute values of $T_n(x)$ in (a, b) is to be obtained in terms of H_{ns}. Let μ_n be the maximum of $| T_n(x) |$ for $a \leqq x \leqq b$, taken on for $x = x_0$, so that $| T_n(x_0) | = \mu_n$. By Theorem 2, $| T_n'(x) | \leqq Cn^2\mu_n$ for $a \leqq x \leqq b$. Hence if $|x - x_0| \leqq 1/(2Cn^2)$, with this particular value of C, the point x remaining in (a, b),

$$| T_n(x) - T_n(x_0) | \leqq \tfrac{1}{2}\mu_n,$$

and $| T_n(x) | \geqq \tfrac{1}{2}\mu_n$. Without impairment of the accuracy of what has been said it can be assumed that $C > 1/(b-a)$, so that $1/(2Cn^2) < \tfrac{1}{2}(b-a)$ for $n \geqq 1$. Then an interval of length $1/(2Cn^2)$ on one side or the other of x_0 is wholly contained in (a, b), and the integrand in the expression defining H_{ns} is greater than or equal to $(\mu_n/2)^s$ throughout such an interval. Consequently

$$H_{ns} \geqq \frac{1}{2Cn^2}\left(\frac{\mu_n}{2}\right)^s, \qquad \mu_n \leqq C_s(n^2 H_{ns})^{1/s}.$$

This may be summarized as follows:

LEMMA 1. *If $T_n(x)$ is a trigonometric sum of the nth order, if*

$$H_{ns} = \int_a^b | T_n(x) |^s dx, \qquad 0 < b - a < 2\pi, \tag{2}$$

and if μ_n is the maximum of $| T_n(x) |$ for $a \leqq x \leqq b$, then

$$\mu_n \leqq C_s(n^2 H_{ns})^{1/s}.$$

---

* Annals of Mathematics, (2), vol. 34 (1933), pp. 799–814.

† For $b - a = 2\pi$ see the paper A, p. 891, Lemma 1. The corresponding statement for $b - a > 2\pi$ follows as a trivial corollary.

Let $\rho(x)$ be a summable function such that $\rho(x) \geqq v > 0$ for $a \leqq x \leqq b$, where v is constant. Then

$$\int_a^b [T_n(x)]^2 dx \leqq \frac{1}{v} \int_a^b \rho(x)[T_n(x)]^2 dx.$$

An immediate consequence of the lemma for $s = 2$ is

THEOREM 7. *If $\rho(x)$ is a summable function having a positive lower bound in (a, b), and if trigonometric sums $T_n(x)$ of the nth order are constructed for successive values of n so that*

$$\int_a^b \rho(x)[T_n(x)]^2 dx \leqq 1, \tag{3}$$

then $|T_n(x)|$ has an upper bound of the order of n for $a \leqq x \leqq b$.

For trigonometric sums normalized over (a, b) with respect to $\rho(x)$ as weight function the integral in (3) is of course equal to 1; the superficially more general statement with the alternative relation of inequality, here and in subsequent theorems, is for the sake of application to the case of sums normalized over a longer interval containing (a, b), but with a weight function not necessarily satisfying the hypothesis throughout the whole of the longer interval.

If $\rho(x)$, without having a positive lower bound, is still non-negative, and is such that $[\rho(x)]^{-r}$ is summable over (a, b) with a positive value of r, let

$$s = 2r/(r+1), \qquad p = 2/s = 1 + (1/r) > 1.$$

The integrand in (2) can be regarded as the product of the factors $[\rho(x)]^{1/p} |T_n(x)|^s$ and $[\rho(x)]^{-1/p}$. By Hölder's inequality, since

$$\{[\rho(x)]^{1/p} |T_n(x)|^s\}^p = \rho(x)[T_n(x)]^2 \text{ and } \{[\rho(x)]^{-1/p}\}^{p/(p-1)} = [\rho(x)]^{-r},$$

$$H_{ns} \leqq \left[\int_a^b \rho(x)[T_n(x)]^2 dx\right]^{1/p} \left[\int_a^b [\rho(x)]^{-r} dx\right]^{(p-1)/p}.$$

If the first integral on the right satisfies (3), it follows that H_{ns} has an upper bound independent of n, and application of the lemma leads to the following conclusion:

THEOREM 8. *If $\rho(x)$ is a non-negative summable function such that $[\rho(x)]^{-r}$ is summable over (a, b), with $r > 0$, and if trigonometric sums $T_n(x)$ of the nth order are constructed for successive values of n so that (3) is satisfied, then $|T_n(x)|$ has an upper bound of the order of $n^{(r+1)/r}$ for $a \leqq x \leqq b$.*

Further information with regard to the order of magnitude of $|T_n(x)|$ in

the interior of the interval can be obtained by the use of a lemma established in the paper A for polynomials.

Consider first the case of a trigonometric sum $C_n(x)$ of the nth order containing only cosines. Let

$$H_{ns} = \int_a^b | C_n(x) |^s dx,$$

with $0 < a < b < \pi$. Let $z = \cos x$, $A = \cos a$, $B = \cos b$. Then $C_n(x)$ is a polynomial $p_n(z)$ of the nth degree in z, and

$$H_{ns} = \int_B^A | p_n(z) |^s \frac{dz}{(1 - z^2)^{1/2}},$$

whence

$$\int_B^A | p_n(z) |^s dz \leqq H_{ns}.$$

From this it follows* that

$$| p_n(z) | \leqq \frac{C_s(nH_{ns})^{1/s}}{[(A - z)(z - B)]^{N/2}}$$

for $B < z < A$, where N is the smallest integer $\geqq 1/s$; if $s \geqq 1$, $N = 1$. Consequently

$$| C_n(x) | \leqq \frac{C_s(nH_{ns})^{1/s}}{[(b - x)(x - a)]^{N/2}}$$

for $a < x < b$, with a new value of C_s.

Let $S_n(x)$ be a sine sum of the nth order, and let

$$H_{ns} = \int_a^b | S_n(x) |^s dx,$$

with $0 < a < b < \pi$ as before. The symbols z, A, B having the same meaning as in the preceding paragraph, $S_n(x) = (1 - z^2)^{1/2} q_{n-1}(z)$, where $q_{n-1}(z)$ is a polynomial in z of degree $n - 1$, and

$$H_{ns} = \int_B^A (1 - z^2)^{(s-1)/2} | q_{n-1}(z) |^s dz, \qquad \int_B^A | q_{n-1}(z) |^s dz \leqq C_s H_{ns}.$$

By the proposition cited for the polynomial $p_n(z)$ above,

$$| q_{n-1}(z) | \leqq \frac{C_s(nH_{ns})^{1/s}}{[(A - z)(z - B)]^{N/2}}, \qquad | S_n(x) | \leqq \frac{C_s(nH_{ns})^{1/s}}{[(b - x)(x - a)]^{N/2}}.$$

* See A, p. 897, Lemma 4.

Let $T_n(x)$ be an arbitrary trigonometric sum of the nth order, let it be supposed again that $0<a<b<\pi$, and let

$$H_{ns} = \int_a^b |\, T_n(x)\,|^s dx + \int_{-b}^{-a} |\, T_n(x)\,|^s dx$$

$$= \int_a^b \{\, |\, T_n(x)\,|^s + |\, T_n(-x)\,|^s \}\, dx.$$

Let $C_n(x) = \frac{1}{2}[T_n(x)+T_n(-x)]$, $S_n(x) = \frac{1}{2}[T_n(x)-T_n(-x)]$, so that $C_n(x)$ is a cosine sum and $S_n(x)$ a sine sum. If X and Y are any two real numbers, and if s is any positive exponent, whether greater than 1, equal to 1, or less than 1, $|\frac{1}{2}(X+Y)|^s$ can not exceed the larger of the quantities $|X|^s$, $|Y|^s$, since the maximum of the function $\Phi(X) \equiv |X|^s$ in any finite interval is taken on at one end or the other of the interval, and consequently

$$|\, \tfrac{1}{2}(X+Y)\,|^s \leqq |\, X\,|^s + |\, Y\,|^s$$

in all cases. In the present connection,

$$|\, C_n(x)\,|^s \leqq |\, T_n(x)\,|^s + |\, T_n(-x)\,|^s, \quad |\, S_n(x)\,|^s \leqq |\, T_n(x)\,|^s + |\, -T_n(-x)\,|^s,$$

so that

$$\int_a^b |\, C_n(x)\,|^s dx \leqq H_{ns}, \qquad \int_a^b |\, S_n(x)\,|^s dx \leqq H_{ns}.$$

Application of the two preceding paragraphs to $C_n(x)$ and $S_n(x)$, and again to $C_n(-x)$ and $S_n(-x)$, gives

$$|\, T_n(x)\,| = |\, C_n(x) + S_n(x)\,| \leqq \frac{C_s(nH_{ns})^{1/s}}{[(b-x)(x-a)]^{N/2}}$$

for $a<x<b$,

$$|\, T_n(x)\,| \leqq \frac{C_s(nH_{ns})^{1/s}}{[(-a-x)(x+b)]^{N/2}}$$

for $-b<x<-a$. These formulas can be adapted by a change of origin to any pair of equal non-contiguous intervals contained within a period.

Let $T_n(x)$ be any trigonometric sum of the nth order, let (a, b) be any interval of length less than 2π, and let

$$H_{ns} = \int_a^b |\, T_n(x)\,|^s dx.$$

Let intervals $(a_1, b_1), \cdots, (a_4, b_4)$ be defined as in connection with the proof

of Theorem 1. The preceding paragraph is to be applied to the first and third of these intervals, and again to the second and fourth; by combination of the results as interpreted for the four quarters of the interval (a, b) a conclusion is reached which can be stated in the following form:

LEMMA 2. *If $T_n(x)$ is a trigonometric sum of the nth order, if*

$$H_{ns} = \int_a^b |T_n(x)|^s dx, \qquad 0 < b - a < 2\pi,$$

and if N is the smallest integer $\geqq 1/s$, then

$$|T_n(x)| \leqq \frac{C_s(nH_{ns})^{1/s}}{[(b-x)(x-a)]^{N/2}}$$

for $a<x<b$.

An immediate consequence of Lemma 2 for $s=2$ is

THEOREM 9. *If $\rho(x)$ is a summable function having a positive lower bound in (a, b), and if trigonometric sums $T_n(x)$ are constructed so that (3) is satisfied, $|T_n(x)|$ has an upper bound of the order of $n^{1/2}$ throughout any closed interval interior to (a, b).*

If $\rho(x)$ does not have a positive lower bound, but does satisfy the hypothesis of Theorem 8, the reasoning with Hölder's inequality in the proof of that theorem, which gave a relation between the integral in (3) and the value of H_{ns} for an appropriate value of s, is applicable again without change; in conjunction with Lemma 2 it gives

THEOREM 10. *If $\rho(x)$ is a non-negative summable function satisfying the hypothesis of Theorem 8, and if trigonometric sums $T_n(x)$ are constructed so that (3) is satisfied, $|T_n(x)|$ has an upper bound of the order of $n^{(r+1)/(2r)}$ throughout any closed interval interior to (a, b).*

Lemma 2 also leads to a new result for polynomials, intermediate in character between Lemmas 3 and 4 of the paper A.

Let $P_n(x)$ be a polynomial of the nth degree, and s a positive exponent, and let

$$h_{ns} = \int_0^1 (1-x)^{-1/2} |P_n(x)|^s dx.$$

Then

$$\int_0^1 (1-x^2)^{-1/2} |P_n(x)|^s dx \leqq h_{ns},$$

since $1+x \geqq 1$ throughout the interval of integration. Let $x=\cos\theta$, so that $P_n(x)=C_n(\theta)$, a cosine sum of the nth order in θ. Then

$$\int_0^{\pi/2} |C_n(\theta)|^s d\theta = \int_0^1 (1-x^2)^{-1/2} |P_n(x)|^s dx \leqq h_{ns},$$

and since $C_n(\theta)$ is an even function of θ

$$\int_{-\pi/2}^{\pi/2} |C_n(\theta)|^s d\theta = 2\int_0^{\pi/2} |C_n(\theta)|^s d\theta \leqq 2h_{ns}.$$

By Lemma 2, therefore,

$$|C_n(\theta)| \leqq \frac{A_s(nh_{ns})^{1/s}}{[(\pi/2)^2-\theta^2]^{N/2}}$$

for $-\pi/2<\theta<\pi/2$, where A_s depends only on s, since the interval is definitely specified. For $0\leqq\theta<\pi/2$ (and so for $0<x\leqq 1$), inasmuch as $(\pi/2)+\theta\geqq\pi/2>1$, while $(\pi/2)-\theta>\cos\theta$,

$$|P_n(x)| = |C_n(\theta)| \leqq \frac{A_s(nh_{ns})^{1/s}}{[(\pi/2)-\theta]^{N/2}} \leqq \frac{A_s(nh_{ns})^{1/s}}{(\cos\theta)^{N/2}} = \frac{A_s(nh_{ns})^{1/s}}{x^{N/2}},$$

with the same A_s throughout. Replacement of x by $(x-a)/(b-a)$, carrying over the interval $(0, 1)$ into an arbitrary interval (a, b), makes it possible to state the conclusion in the following form:

LEMMA 3. *If $P_n(x)$ is a polynomial of the nth degree, if*

$$h_{ns} = \int_a^b (b-x)^{-1/2} |P_n(x)|^s dx,$$

and if N is the smallest integer $\geqq 1/s$, then

$$|P_n(x)| \leqq \frac{C_s(nh_{ns})^{1/s}}{(x-a)^{N/2}}$$

for $a<x\leqq b$.

Similarly, if

$$h_{ns} = \int_a^b (x-a)^{-1/2} |P_n(x)|^s dx,$$

then

$$|P_n(x)| \leqq \frac{C_s(nh_{ns})^{1/s}}{(b-x)^{N/2}}$$

for $a\leqq x<b$.

The second statement follows from the first, without re-examination of the proof in detail, on replacement of x, a, b by $-x$, $-b$, $-a$.

Lemma 3 has the following consequences, one immediate, the other obtained by the use of Hölder's inequality with $[\rho(x)]^{1/p}|P_n(x)|^s$ and $(b-x)^{-1/2}[\rho(x)]^{-1/p}$ as factors under the integral sign, p and s being related to r in the same way as before; each theorem has an alternative statement with the ends of the interval interchanged:

THEOREM 11. *If $\rho(x)$ is a summable function such that $(b-x)^{1/2}\rho(x)$ has a positive lower bound for $a \leqq x < b$, and if polynomials $P_n(x)$ are defined for successive values of n so that*

$$\int_a^b \rho(x)[P_n(x)]^2 dx \leqq 1, \tag{4}$$

then $|P_n(x)|$ has an upper bound of the order of $n^{1/2}$ throughout any interval $a+\delta \leqq x \leqq b$, $\delta > 0$.

THEOREM 12. *If $\rho(x)$ is a non-negative summable function such that $(b-x)^{-(r+1)/2}[\rho(x)]^{-r}$ is summable over (a, b), with $r>0$, and if polynomials $P_n(x)$ are defined satisfying (4), then $|P_n(x)|$ has an upper bound of the order of $n^{(r+1)/(2r)}$ throughout any interval $a+\delta \leqq x \leqq b$, $\delta > 0$.*

From Lemma 3 it is possible to proceed to a trigonometric analogue of Lemma 3 of the paper A.

Let $C_n(x)$ be a cosine sum of the nth order, and let

$$h_{ns} = \int_a^b (b-x)^{-1/2}|C_n(x)|^s dx, \qquad 0 < a < b < \pi.$$

With the notation used in the proof of Lemma 2,

$$\int_{B'}^{A} (z-B)^{-1/2}|p_n(z)|^s dz \leqq C h_{ns},$$

whence by Lemma 3

$$|C_n(x)| = |p_n(z)| \leqq \frac{C_s(nh_{ns})^{1/s}}{(A-z)^{N/2}} \leqq \frac{C_s(nh_{ns})^{1/s}}{(x-a)^{N/2}}$$

in (a, b). A similar argument, with obvious modifications, applies to a sum of sines.

Let $T_n(x)$ be an arbitrary trigonometric sum of the nth order, and with the assumption still that $0<a<b<\pi$ let

$$H'_{ns} = \int_a^b (b^2 - x^2)^{-1/2} \left| T_n(x) \right|^s dx + \int_{-b}^{-a} (b^2 - x^2)^{-1/2} \left| T_n(x) \right|^s dx$$

$$= \int_a^b (b^2 - x^2)^{-1/2} (\left| T_n(x) \right|^s + \left| T_n(-x) \right|^s) dx.$$

Then

$$\int_a^b (b - x)^{-1/2} (\left| T_n(x) \right|^s + \left| T_n(-x) \right|^s) dx \leqq CH'_{ns}.$$

If $S_n(x)$ and $C_n(x)$ are the odd and even parts of $T_n(x)$,

$$\int_a^b (b - x)^{-1/2} \left| C_n(x) \right|^s dx \leqq CH'_{ns}, \qquad \int_a^b (b - x)^{-1/2} \left| S_n(x) \right|^s dx \leqq CH'_{ns},$$

some of the steps in the proof of Lemma 2 being applicable once more, and

$$\left| C_n(x) \right| \leqq \frac{C_s(nH'_{ns})^{1/s}}{(x - a)^{N/2}}, \qquad \left| S_n(x) \right| \leqq \frac{C_s(nH'_{ns})^{1/s}}{(x - a)^{N/2}}$$

for $a < x \leqq b$. So $\left| T_n(x) \right|$ has an upper bound of the same form in this interval. Replacement of x by $-x$ in the even functions $\left| C_n(x) \right|$ and $\left| S_n(x) \right|$ gives for each of these functions, and hence for $\left| T_n(x) \right|$, an upper bound of the form $C_s(nH'_{ns})^{1/s}/(-x-a)^{N/2}$ in the interval $-b \leqq x < -a$. Consequently

$$\left| T_n(x) \right| \leqq \frac{C_s(nH'_{ns})^{1/s}}{(x^2 - a^2)^{N/2}}$$

in both intervals.

By a change of origin, if (a_1, b_1) and (a_2, b_2) are any two equal non-contiguous intervals contained within a period, and if

$$H'_{ns} = \int_{a_1}^{b_1} [(b_2 - x)(x - a_1)]^{-1/2} \left| T_n(x) \right|^s dx + \int_{a_2}^{b_2} [(b_2 - x)(x - a_1)]^{-1/2} \left| T_n(x) \right|^s dx,$$

the intervals (a_1, b_1) and (a_2, b_2) taking the place of $(-b, -a)$ and (a, b) respectively, then

$$\left| T_n(x) \right| \leqq \frac{C_s(nH'_{ns})^{1/s}}{[(x - a_2)(x - b_1)]^{N/2}} \tag{5}$$

for $a_1 \leqq x < b_1$ and for $a_2 < x \leqq b_2$.

The desired conclusion with regard to an arbitrary trigonometric sum $T_n(x)$ in a single arbitrary interval (a, b) (of length less than 2π) is to be

derived now by the use of *three* overlapping intervals covering (a, b). To avoid another definition of symbols the notation used at the corresponding stage in the proofs of Theorem 1 and Lemma 2 may be employed again, the three intervals for the present argument being taken as (a, b_1), (a_2, b_3), and (a_4, b). Let

$$H'_{ns} = \int_a^b [(b - x)(x - a)]^{-1/2} \mid T_n(x) \mid^s dx.$$

Then

$$\int_a^{b_1} [(b - x)(x - a)]^{-1/2} \mid T_n(x) \mid^s dx$$

$$+ \int_{a_4}^{b} [(b - x)(x - a)]^{-1/2} \mid T_n(x) \mid^s dx \leqq H'_{ns},$$

and by (5)

$$\mid T_n(x) \mid \leqq \frac{C_s(nH'_{ns})^{1/s}}{[(x - a_4)(x - b_1)]^{N/2}}$$

in (a, b_1) and in (a_4, b). If x is restricted to the closed intervals (a, c_1) and (c_3, b), in which the last denominator has a positive lower bound, $\mid T_n(x) \mid \leqq C_s(nH'_{ns})^{1/s}$. On the other hand,

$$\int_{a_2}^{b_3} \mid T_n(x) \mid^s dx \leqq CH'_{ns},$$

so that Lemma 2 gives a relation of the form $\mid T_n(x) \mid \leqq C_s(nH_{ns}')^{1/s}$ for the interval $c_1 \leqq x \leqq c_3$, interior to (a_2, b_3). For the whole interval (a, b) a conclusion is thus obtained which can be expressed as

LEMMA 4. *If $T_n(x)$ is a trigonometric sum of the nth order and if*

$$H'_{ns} = \int_a^b [(b - x)(x - a)]^{-1/2} \mid T_n(x) \mid^s dx, \qquad 0 < b - a < 2\pi,$$

then $\mid T_n(x) \mid \leqq C_s(nH'_{ns})^{1/s}$ for $a \leqq x \leqq b$.

This lemma leads to inferences corresponding to those associated with the earlier lemmas:

THEOREM 13. *If $\rho(x)$ is a summable function such that $[(b-x)(x-a)]^{1/2}\rho(x)$ has a positive lower bound for $a<x<b$, and if trigonometric sums $T_n(x)$ are constructed satisfying (3), then $\mid T_n(x) \mid$ has an upper bound of the order of $n^{1/2}$ for $a \leqq x \leqq b$.*

THEOREM 14. *If $\rho(x)$ is a non-negative summable function such that*

$$[(b-x)(x-a)]^{-(r+1)/2}[\rho(x)]^{-r}$$

is summable over (a, b), with $r>0$, and if trigonometric sums $T_n(x)$ are constructed satisfying (3), then $|T_n(x)|$ has an upper bound of the order of $n^{(r+1)/(2r)}$ for $a \leqq x \leqq b$.

5. **Convergence theorems.** In this section there will be presented a sequence of applications of the preceding ideas to the theory of convergence of processes of closest approximation. A process of reasoning which has been employed elsewhere,* modified by the use of Theorem 2 in place of the ordinary form of Bernstein's or Markoff's theorem, establishes the following proposition:

LEMMA 5. *If $f(x)$ is a continuous function for $a \leqq x \leqq b$, $T_n(x)$ a trigonometric sum of the nth order, and*

$$G_{ns} = \int_a^b |f(x) - T_n(x)|^s dx,$$

and if there exists a trigonometric sum $t_n(x)$ of the nth order such that

$$|f(x) - t_n(x)| \leqq \epsilon_n$$

throughout the interval, then

$$|f(x) - T_n(x)| \leqq C_s(n^2 G_{ns})^{1/s} + 5\epsilon_n$$

for $a \leqq x \leqq b$.

Now let it be supposed that $f(x)$ is defined and continuous for $\alpha \leqq x \leqq \beta$, where $\alpha \leqq a < b \leqq \beta$ and $\beta - \alpha \leqq 2\pi$. (A more general hypothesis would be admissible for the moment, but unprofitable.) If $\beta = \alpha + 2\pi$ it will be assumed that $f(\beta) = f(\alpha)$, so that $f(x)$ can be thought of as continuous and periodic. Let $\rho(x)$ be a non-negative summable function in (α, β), having a positive constant v as lower bound for $a \leqq x \leqq b$. Let m be a given positive exponent, and let $T_n(x)$ be chosen among all trigonometric sums of the nth order so as to minimize the integral

$$\int_\alpha^\beta \rho(x) |f(x) - T_n(x)|^m dx. \tag{6}$$

The question of the existence of a minimizing sum, as well as of its uniqueness or non-uniqueness, has been treated elsewhere† in a manner adequate

* See, for example, the paper A, p. 899; also cf. the proof of Lemma 1 above.

† See these Transactions, vol. 26, loc. cit., pp. 133–139.

for present purposes; if $\beta-\alpha<2\pi$ the integral can be regarded nevertheless as extended over an entire period, with a weight function $\rho(x)$ vanishing outside (α, β). Let the minimum value of the integral, corresponding to the specific sum $T_n(x)$ under consideration, be γ_n. Then

$$G_{nm} = \int_a^b | f(x) - T_n(x) |^m dx \leqq \frac{1}{v} \int_a^b \rho(x) | f(x) - T_n(x) |^m dx \leqq \frac{\gamma_n}{v},$$

so that by Lemma 5

$$| f(x) - T_n(x) | \leqq C_m(n^2\gamma_n/v)^{1/m} + 5\epsilon_n$$

in (a, b), if ϵ_n has the meaning assigned to it above. Let it be assumed now that $f(x)$ can be approximately represented with an error not exceeding ϵ_n by a trigonometric sum of the nth order for $\alpha \leqq x \leqq \beta$, and not merely in (a, b), if the former interval is more extensive. Then by the minimizing property of $T_n(x)$

$$\gamma_n \leqq R\epsilon_n^m, \qquad R = \int_\alpha^\beta \rho(x)dx.$$

It follows that

$$| f(x) - T_n(x) | \leqq kn^{2/m}\epsilon_n$$

for $a \leqq x \leqq b$, with k independent of x and n. This result will be stated as a theorem.*

THEOREM 15. *If $\rho(x)$ is a non-negative summable function over (α, β) having a positive lower bound in (a, b), $\alpha \leqq a < b \leqq \beta$, if trigonometric sums $T_n(x)$ of the nth order are constructed for successive values of n to minimize the integral* (6), *and if trigonometric sums $t_n(x)$ of the nth order exist for each n so that*

$$| f(x) - t_n(x) | \leqq \epsilon_n$$

for $\alpha \leqq x \leqq \beta$, there is a constant k, independent of x and n, such that

$$| f(x) - T_n(x) | \leqq kn^{2/m}\epsilon_n$$

for $a \leqq x \leqq b$.

A sufficient condition for uniform convergence of $T_n(x)$ toward $f(x)$ throughout (a, b) is that $\lim_{n\to\infty} n^{2/m}\epsilon_n = 0$. Sums $t_n(x)$ meeting this requirement will exist if the modulus of continuity of $f(x)$, or in case of need, according to the value of m, the modulus of continuity of one of its derivatives, is suitably restricted.† If $m=2$, for example, it is sufficient that $f(x)$ have

* See also these Transactions, vol. 26, loc. cit., pp. 145–150.

† See, for example, Colloquium, Chapter I.

a continuous derivative for $\alpha \leqq x \leqq \beta$. Although the general theorems referred to for the existence of approximating sums are stated for the approximate representation of a periodic function over an entire period they are nevertheless available for use here, since if the interval (α, β) in which $f(x)$ is originally defined is of length less than 2π the definition can in each case be extended through the rest of a period so that the requisite properties are preserved.*

The remaining theorems to be presented below will be stated in a form corresponding to that of Theorem 15, the explicit interpretation of the results in terms of properties of continuity of $f(x)$ being omitted.

If $\rho(x)$ does not have a positive lower bound in (a, b), but is such that $[\rho(x)]^{-r}$ is summable over (a, b), $r>0$, let

$$s = mr/(r+1), \qquad p = m/s = 1 + (1/r) > 1.$$

By Hölder's inequality G_{ns}, written in the form

$$\int_a^b [\rho(x)]^{-1/p} \cdot [\rho(x)]^{1/p} \, | f(x) - T_n(x) |^s \, dx,$$

does not exceed the product of

$$\left[\int_a^b \rho(x) \, | f(x) - T_n(x) |^m dx \right]^{1/p}$$

by a factor independent of n. The integral last written down does not exceed γ_n, which is the integral of the same integrand from α to β. As $\gamma_n \leqq R\epsilon_n^m$, G_{ns} is not greater than a constant multiple of ${\epsilon_n}^{m/p} = {\epsilon_n}^s$, and $|f(x) - T_n(x)|$ has by Lemma 5 an upper bound of the order of $n^{2/s}\epsilon_n$. This gives

THEOREM 16. *If $\rho(x)$ is a non-negative summable function over (α, β) such that $[\rho(x)]^{-r}$ is summable over (a, b), $\alpha \leqq a < b \leqq \beta$, $r>0$, if trigonometric sums $T_n(x)$ are defined minimizing* (6), *and if ϵ_n has the same meaning as in Theorem* 15, *there is a constant k, independent of x and n, such that*

$$| f(x) - T_n(x) | \leqq k n^{2(r+1)/(mr)} \epsilon_n$$

for $a \leqq x \leqq b$.

The lemma to be presented next corresponds to Lemma 8 of the paper A, and is to be deduced as a direct consequence of it.

Let $f(x)$ be continuous for $a \leqq x \leqq b$, with the understanding for the present that $0 < a < b < \pi$. Let $C_n(x)$ be a cosine sum of the nth order, and let

* If $f(x)$ has a modulus of continuity $\omega(\delta)$ for $\alpha \leqq x \leqq \beta$ and is defined as linear for $\beta \leqq x \leqq \alpha + 2\pi$, it has a modulus of continuity not exceeding a constant multiple of $\omega(\delta)$ for $\alpha \leqq x \leqq \alpha + 2\pi$; see Colloquium, pp. 52–53. The extension to derivatives offers no difficulty.

$$G_{ns} = \int_a^b | f(x) - C_n(x) |^s dx.$$

Let $z = \cos x$, $A = \cos a$, $B = \cos b$, $f(x) = \phi(z)$. Let $C_n(x)$ as a polynomial in $\cos x$ be denoted by $p_n(z)$. Then

$$G_{ns} = \int_B^A | \phi(z) - p_n(z) |^s (1 - z^2)^{-1/2} dz, \qquad \int_B^A | \phi(z) - p_n(z) |^s dz \leqq G_{ns}.$$

If there is a cosine sum of the nth order differing from $f(x)$ by not more than ϵ_n for $a \leqq x \leqq b$, this is a polynomial of the nth degree in z differing from $\phi(z)$ by not more than ϵ_n for $B \leqq z \leqq A$, and by the lemma* referred to in the preceding paragraph

$$| f(x) - C_n(x) | = | \phi(z) - p_n(z) |$$
$$\leqq \frac{C_s[(nG_{ns})^{1/s} + \epsilon_n]}{[(A - z)(z - B)]^{N/2}} \leqq \frac{C_s[(nG_{ns})^{1/s} + \epsilon_n]}{[(b - x)(x - a)]^{N/2}}$$

for $a < x < b$, N being the smallest integer $\geqq 1/s$.

In carrying through a similar argument with sines in place of cosines some changes are to be noted in details. Let $f(x)$ be as before, let $S_n(x)$ be a sine sum of the nth order, and let G_{ns} now denote the integral

$$G_{ns} = \int_a^b | f(x) - S_n(x) |^s dx, \qquad 0 < a < b < \pi.$$

Let the symbols z, A, B, $\phi(z)$ be interpreted as above. The sum $S_n(x)$ has the form $(1-z^2)^{1/2} q_{n-1}(z)$, where $q_{n-1}(z)$ is a polynomial of degree $n-1$. Let

$$\psi(z) = \frac{\phi(z)}{(1 - z^2)^{1/2}} = \frac{f(x)}{\sin x}.$$

Then

$$\int_B^A | \psi(z) - q_{n-1}(z) |^s dz \leqq C_s G_{ns}.$$

If there is a sine sum $s_n(x)$ of the nth order differing from $f(x)$ by not more than ϵ_n for $a \leqq x \leqq b$ then $s_n(x)/\sin x$ is a polynomial of degree $n-1$ in z differing from $\psi(z)$ by not more than $C\epsilon_n$ for $B \leqq z \leqq A$. Another application of Lemma 8 of the paper A gives for $|f(x) - S_n(x)|$ an upper bound of the form obtained for $|f(x) - C_n(x)|$ at the end of the last paragraph.

Let $T_n(x)$ be an arbitrary trigonometric sum of the nth order, and with another change in the meaning of G_{ns} let

* See A, p. 905.

$$G_{ns} = \int_a^b |f(x) - T_n(x)|^s dx + \int_{-b}^{-a} |f(x) - T_n(x)|^s dx$$

$$= \int_a^b \{|f(x) - T_n(x)|^s + |f(-x) - T_n(-x)|^s\} dx, \qquad 0 < a < b < \pi.$$

Let

$$u(x) = \tfrac{1}{2}[f(x) + f(-x)], \qquad v(x) = \tfrac{1}{2}[f(x) - f(-x)],$$
$$C_n(x) = \tfrac{1}{2}[T_n(x) + T_n(-x)], \qquad S_n(x) = \tfrac{1}{2}[T_n(x) - T_n(-x)].$$

By an argument used in connection with the proof of Lemma 2,

$$\int_a^b |u(x) - C_n(x)|^s dx \leqq G_{ns}.$$

Let it be supposed that there is a trigonometric sum $t_n(x)$ such that ϵ_n is an upper bound for $|f(x)-t_n(x)|$ throughout both of the intervals (a, b), $(-b, -a)$. Then if $c_n(x)$ is the cosine sum $\frac{1}{2}[t_n(x)+t_n(-x)]$,

$$|u(x) - c_n(x)| \leqq \epsilon_n$$

for $a \leqq x \leqq b$. So the second paragraph preceding gives an upper bound for $|u(x)-C_n(x)|$ in the interval $a<x<b$. Similarly, $|v(x)-S_n(x)|$ has an upper bound of the same form over the same range. As substitution of $-x$ for x leads to corresponding results for the interval $(-b, -a)$, it is found that

$$|f(x) - T_n(x)| \leqq \frac{C_s[(nG_{ns})^{1/s} + \epsilon_n]}{[(-a - x)(x + b)]^{N/2}},$$

$$|f(x) - T_n(x)| \leqq \frac{C_s[(nG_{ns})^{1/s} + \epsilon_n]}{[(b - x)(x - a)]^{N/2}}$$

for $-b<x<-a$ and for $a<x<b$ respectively. These formulas can be adapted immediately to the case of a pair of equal intervals unsymmetrically situated with respect to the origin. The transition to the case of a single arbitrary interval of length less than 2π is then effected by the use of four overlapping subintervals as in the proofs of Theorem 1 and Lemma 2, the conclusion being as follows:

LEMMA 6. *Under the hypotheses of Lemma* 5

$$|f(x) - T_n(x)| \leqq \frac{C_s[(nG_{ns})^{1/s} + \epsilon_n]}{[(b - x)(x - a)]^{N/2}}$$

for $a<x<b$, N being the smallest integer $\geqq 1/s$.

If the reasoning which led from Lemma 5 to Theorem 15 is modified by

the use of Lemma 6 in place of Lemma 5 a result is obtained, the most essential content of which can be expressed thus:

THEOREM 17. *Under the hypotheses of Theorem* 15, $|f(x)-T_n(x)|$ *has an upper bound of the order of* $n^{1/m}\epsilon_n$ *throughout any closed interval interior to* (a, b).

This is an appreciable improvement over an earlier result* of similar character, to the extent that the exponent $(1/m)+\eta$ of the earlier work is replaced here by $1/m$.

Lemma 6 in combination with Hölder's inequality (as applied in the proof of Theorem 16) gives

THEOREM 18. *Under the hypotheses of Theorem* 16, $|f(x)-T_n(x)|$ *has an upper bound of the order of* $n^{(r+1)/(mr)}\epsilon_n$ *throughout any closed interval interior to* (a, b).

Lemma 6 also has a bearing on the theory of polynomial approximation, leading to a lemma which is related to it as Lemma 3 is related to Lemma 2.

Let $f(x)$ be a continuous function for $0 \leqq x \leqq 1$, let $P_n(x)$ be a polynomial of the nth degree, and let

$$g_{ns} = \int_0^1 (1-x)^{-1/2} |f(x) - P_n(x)|^s dx.$$

If $x=\cos\theta$, $f(x)=\psi(\theta)$, $P_n(x)=C_n(\theta)$, application of Lemma 6 in the manner suggested by the proof of Lemma 3 gives in terms of g_{ns} an upper bound for $|\psi(\theta)-C_n(\theta)|$ in the interval $-\pi/2<\theta<\pi/2$, and so for $|f(x)-P_n(x)|$ in the interval $0<x\leqq 1$. Interpreted for an arbitrary interval by a linear change of variable, the conclusion may be stated as

LEMMA 7. *If* $f(x)$ *is a continuous function for* $a \leqq x \leqq b$, $P_n(x)$ *a polynomial of the nth degree, and*

$$g_{ns} = \int_a^b (b-x)^{-1/2} |f(x) - P_n(x)|^s dx,$$

and if there exists a polynomial $p_n(x)$ *of the nth degree such that*

$$|f(x) - p_n(x)| \leqq \epsilon_n$$

throughout the interval, then

$$|f(x) - P_n(x)| \leqq \frac{C_s[(ng_{ns})^{1/s} + \epsilon_n]}{(x-a)^{N/2}}$$

for $a<x\leqq b$, N *being the smallest integer* $\geqq 1/s$.

* These Transactions, vol. 26, loc. cit., p. 153. For comparison see also the paper A, p. 906, Theorem 15 and context.

If

$$g_{ns} = \int_a^b (x-a)^{-1/2} \left| f(x) - P_n(x) \right|^s dx$$

then

$$\left| f(x) - P_n(x) \right| \leqq \frac{C_s[(ng_{ns})^{1/s} + \epsilon_n]}{(b-x)^{N/2}}$$

for $a \leqq x < b$.

The second statement is an immediate corollary of the first.

Theorems on polynomials of closest approximation resulting from Lemma 7 are as follows:

THEOREM 19. *If* $\rho(x)$ *is a non-negative summable function over* (α, β) *such that* $(b-x)^{1/2}\rho(x)$ *has a positive lower bound in* (a, b), $\alpha \leqq a < b \leqq \beta$, *if polynomials* $P_n(x)$ *of the nth degree are constructed for successive values of n to minimize the integral*

$$\int_\alpha^\beta \rho(x) \left| f(x) - P_n(x) \right|^m dx, \tag{7}$$

and if polynomials $p_n(x)$ *of the nth degree exist for each n so that*

$$\left| f(x) - p_n(x) \right| \leqq \epsilon_n$$

for $\alpha \leqq x \leqq \beta$, *then* $|f(x) - P_n(x)|$ *has an upper bound of the order of* $n^{1/m}\epsilon_n$ *throughout any interval* $a + \delta \leqq x \leqq b$, $\delta > 0$.

THEOREM 20. *If* $\rho(x)$ *is a non-negative summable function over* (α, β) *such that* $(b-x)^{-(r+1)/2}[\rho(x)]^{-r}$ *is summable over* (a, b), $\alpha \leqq a < b \leqq \beta$, $r > 0$, *if polynomials* $P_n(x)$ *are defined minimizing* (7), *and if* ϵ_n *has the same meaning as in Theorem* 19, $|f(x) - P_n(x)|$ *has an upper bound of the order of* $n^{(r+1)/(mr)}\epsilon_n$ *throughout any interval* $a + \delta \leqq x \leqq b$, $\delta > 0$.

Each theorem has an alternative statement with $x - a$ in place of $b - x$ and with $(a, b - \delta)$ in place of $(a + \delta, b)$.

The final lemma of the current sequence is to correspond to Lemma 7 of the paper A. It is derived from Lemma 7 of the present paper, with incidental use of Lemma 6, by adaptation of the procedure which led from Lemmas 2 and 3 to Lemma 4, the last stage of the argument involving consideration of three overlapping intervals. Details of the proof being omitted, the statement is this:

LEMMA 8. *If* $f(x)$ *is continuous for* $a \leqq x \leqq b$, $T_n(x)$ *is a trigonometric sum of the nth order, and*

$$G'_{ns} = \int_a^b [(b - x)(x - a)]^{-1/2} | f(x) - T_n(x) |^s dx,$$

and if there is a trigonometric sum $t_n(x)$ of the nth order such that

$$| f(x) - t_n(x) | \leqq \epsilon_n$$

throughout the interval, then

$$| f(x) - T_n(x) | \leqq C_s[(nG'_{ns})^{1/s} + \epsilon_n]$$

for $a \leqq x \leqq b$.

The following are the corresponding theorems on closest approximation:

THEOREM 21. *If $\rho(x)$ is a non-negative summable function over (α, β) such that $[(b-x)(x-a)]^{1/2}\rho(x)$ has a positive lower bound in (a, b), $\alpha \leqq a < b \leqq \beta$, if trigonometric sums $T_n(x)$ are defined minimizing* (6), *and if ϵ_n has the same meaning as in Theorem* 15, *$|f(x)-T_n(x)|$ has an upper bound of the order of $n^{1/m}\epsilon_n$ for $a \leqq x \leqq b$.*

THEOREM 22. *If $\rho(x)$ is a non-negative summable function over (α, β) such that*

$$[(b - x)(x - a)]^{-(r+1)/2}[\rho(x)]^{-r}$$

is summable over (a, b), $\alpha \leqq a < b \leqq \beta$, $r > 0$, if trigonometric sums $T_n(x)$ are defined minimizing (6), *and if ϵ_n has the same meaning as in Theorem* 15, *$|f(x)-T_n(x)|$ has an upper bound of the order of $n^{(r+1)/(mr)}\epsilon_n$ for $a \leqq x \leqq b$.*

THE UNIVERSITY OF MINNESOTA,
MINNEAPOLIS, MINN.

SINGULAR QUADRATIC FUNCTIONALS*

BY

MARSTON MORSE AND WALTER LEIGHTON

Introduction. Singular quadratic functions of the type (1.1) have been briefly investigated by Kemble [1] in connection with Quantum Mechanics. The results of Kemble are of a relatively restricted character. On the other hand, Hardy, Littlewood, and Pólya,† in Chapter VII of [1] have studied special examples of these functionals employing special methods. The developments of this paper apparently give a systematic approach to the problem of minimizing singular quadratic functionals of the type (1.1). In particular, the results obtained include and generalize Theorems (254) and (253) of H.L.P. See Examples 9.2 and 12.1 of this paper.

The authors have admitted various classes of comparison curves. The results show in striking fashion how the existence of the minimum depends upon the classes of curves admitted.

The problem requires a remodeling of the conjugate point theory and an introduction of a new condition called the singularity condition. Lebesgue integrals or their extensions are used throughout. The results of this paper will be applied to extend the theory of characteristic roots and solutions of the related boundary problems.

I. One singular end point

1. The functional. We consider a function

$$f(x, y, y') = r(x)y'^2 + 2q(x)yy' + p(x)y^2, \tag{1.1}$$

where r, q, and p are single-valued, continuous functions of the real variable x on the open interval‡ $(0, d)$ and r is positive. We choose a constant b on the interval $(0, d)$ and consider the functional

$$J(y)\Big|_e^b = \int_e^b f(x, y, y')dx \qquad (0 < e < b < d). \tag{1.2}$$

---

* Presented to the Society, April 11, 1936; received by the editors December 7, 1935.

† We shall refer to these authors as H.L.P.

‡ In designating intervals it will be convenient to use the conventions:

$[a, b]$ means the interval $a \leqq x \leqq b$,
$(a, b]$ means the interval $a < x \leqq b$,
(a, b) means the interval $a < x < b$,
$[a, b)$ means the interval $a \leqq x < b$.

We shall term $y(x)$ and the curve $y=y(x)$ A-admissible on $[0, b]$ if

1. $y(x)$ is continuous on $[0, b]$ and $y(0)=y(b)=0$;

2. $y(x)$ is absolutely continuous and $y'^2(x)$ is in L on each closed subinterval of $(0, b]$.

On an A-admissible curve y'^2 will not in general be in L on $[0, b]$. For example,

$$y = x^{1/2} \cos x$$

is A-admissible on $[0, \pi]$ but y'^2 is not in L on this interval. Our A-admissible functions $y(x)$ are thus less restricted than the admissible functions of H.L.P.*

Observe that the segment $[0, b]$ of the x-axis is A-admissible and that on this segment $J=0$. We shall seek conditions under which

$$\liminf_{x=0} \int_x^b f(x, y, y')dx \geqq 0 \qquad (x > 0). \dagger \tag{1.3}$$

If (1.3) holds for a given class of curves $y=y(x)$ we shall say that $[0, b]$ affords a *minimum limit* to J among curves of the given class. This minimum limit will be termed *proper* if the equality sign in (1.3) holds at most when $y \equiv 0$.

As we shall see in later sections there are functionals J and intervals $[0, b]$ such that $[0, b]$ does not afford a minimum limit to J among A-admissible curves, but such that $[0, b]$ does afford a minimum limit to J if the class of A-admissible curves be further restricted. This fact is particularly important in the cases where the Euler equations have a regular singular point at $x=0$.

2. **Conjugate points.** We seek the analogue of the Jacobi condition. Classical definitions however of the point conjugate to a given point P break down in case P is the point $(0, 0)$ and the theory of conjugate points requires additional development.

We begin with the Euler equation (written E.E.)

$$\frac{d}{dx}(ry' + qy) - (qy' + py) = 0. \tag{2.1}$$

In general the E.E. is singular at $x=0$. If $u(x)$ is a solution of the E.E. of class C^2 on $(0, d)$, we shall term the curve $y=u(x)$ an *extremal*.

We come to the definition of the first conjugate point of $x=0$. Let $x=a$

* Hardy, Littlewood, and Pólya [1], Chapter VII.

† This restriction will be understood in the future.

be a point on $(0, b)$ and $x(a)$ the first conjugate point of $x=a$ following $x=a$ on the interval (a, d), if this conjugate point exists. If $x(a)$ exists for a value $a=a_1$, it also exists for $0<a<a_1$ and decreases with a; the *first conjugate point* $x=\alpha$ of $x=0$ on $[0, d)$ is then defined as the limit of $x(a)$ as a tends to zero. If $x(a)$ exists for no value of a on $(0, b)$, $x=0$ will be said to have no conjugate point on $[0, d)$. The first conjugate point of $x=0$ may coincide with $x=0$ as we shall see in §7.

We shall prove the following theorem.

THEOREM 2.1. *If the first conjugate point $x=\alpha$ of $x=0$ exists and $\alpha\neq 0$, there exists a solution $u(x)$ of the* E.E. *such that $u(\alpha)=0$ and $u(x)>0$ on $(0, \alpha)$. If there is no point on $[0, d)$ conjugate to $x=0$, there exists a solution $v(x)>0$ on $(0, d)$.*

It is sufficient to take $u(x)$ in the theorem as the solution for which

$$u(\alpha) = 0, \qquad u'(\alpha) = -1.$$

To determine a solution $v(x)$ which satisfies the theorem let a and b be two numbers such that $0<a<b<d$ and let $y(x, a)$ be a solution such that

$$y(a, a) = 0, \qquad y(b, a) = 1.$$

Observe that $y'(b, a)$ decreases as a decreases and tends to a limit λ as a tends to zero. The solution $v(x)$ such that

$$v(b) = 1, \qquad v'(b) = \lambda,$$

satisfies the theorem.

We seek to determine the first conjugate point of $x=0$ in terms of solutions of the E.E. characterized neighboring $x=0$. We begin with the following lemma.

LEMMA 2.1. *If there exist linearly independent solutions $u(x)$, $v(x)$ of the* E.E. *for which*

$$\lim_{x=0} \frac{u(x)}{v(x)} = k, \tag{2.2}$$

where k is a finite constant, the solution

$$w(x) = u(x) - kv(x) \tag{2.3}$$

has the following properties.

If $y(x)$ is a solution of the E.E. *which is linearly independent of $w(x)$,*

$$\lim_{x=0} \frac{w(x)}{y(x)} = 0 \tag{2.4}$$

and

$$\frac{w'(x)}{w(x)} - \frac{y'(x)}{y(x)} > 0 \tag{2.5}$$

for sufficiently small positive values of x.

Observe that

$$\lim_{x=0} \frac{w(x)}{v(x)} = 0, \tag{2.6}$$

so that $w(x)$ and $v(x)$ are linearly independent. Any solution $y(x)$ linearly independent of $w(x)$ can be represented in the form

$$y(x) = c_1 w(x) + c_2 v(x) \qquad (c_2 \neq 0),$$

and it follows from (2.6) that (2.4) holds as stated.

To establish (2.5) write the left member of (2.5) in the form

$$\frac{w'}{w} - \frac{c_1 w' + c_2 v'}{c_1 w + c_2 v} = \frac{c_2(w'v - wv')}{wy} = \frac{1}{2}\left(\frac{c_2 v}{y}\right)\left(\frac{v}{w}\right)^2 \frac{d}{dx}\left(\frac{w}{v}\right)^2. \tag{2.7}$$

The first parenthesis in (2.7) tends to one as x tends to zero. Moreover for x sufficiently small and positive

$$\frac{d}{dx}\left(\frac{w}{v}\right)^2 = 2\left(\frac{w}{v}\right)\frac{w'v - wv'}{v^2} \neq 0,$$

since the solutions $w(x)$ and $v(x)$ are independent. It follows from (2.6) that

$$\frac{d}{dx}\left(\frac{w}{v}\right)^2 > 0$$

for x sufficiently small and positive.

The proof of the lemma is complete.

We continue with the two following lemmas.

LEMMA 2.2. *If the solution* $w(x)$ *in the preceding lemma has a first positive zero* $x = c$, *any solution* $y(x)$ *which is independent of* $w(x)$ *vanishes on* $(0, c)$.

The solutions $w(x)$ and $y(x)$ satisfy (2.5) for x sufficiently small and positive. The lemma follows at once from a Sturm comparison theorem.

LEMMA 2.3. *If* $x = 0$ *is not its own first conjugate point, the ratio* u/v *of any two linearly independent solutions tends to a finite limit or becomes infinite as* x *tends to zero.*

The lemma follows at once upon computing the derivative of u/v.

The preceding lemmas lead to the following theorem.

THEOREM 2.2. *If $x=0$ is not its own first conjugate point, there exists a solution $w(x)\not\equiv 0$ such that for each solution $y(x)$ independent of $w(x)$*

$$\lim_{x=0} \frac{w(x)}{y(x)} = 0. \tag{2.8}$$

Any solution with this property of $w(x)$ is linearly dependent on $w(x)$.

The first conjugate point of $x=0$, if it exists, is the first positive zero of $w(x)$, and the first positive zero of $w(x)$, if it exists, is the first conjugate point of $x=0$.

We see from Lemma 2.3 that there exist two linearly independent solutions u and v such that the ratio u/v tends to a finite limit, as x tends to zero. It follows from Lemma 2.1 that there exists a solution $w(x)$ for which (2.8) holds for each solution $y(x)$ independent of $w(x)$, and it follows from (2.8) that the only solutions with this property of $w(x)$ are dependent on $w(x)$.

It remains to prove the two concluding statements in the theorem.

Suppose that $x=\alpha$ is a first conjugate point on $(0, b]$ of $x=0$. It follows from the definition of a conjugate point and from the Sturm Separation Theorem* that $w(x)$ vanishes on $(0, \alpha]$. But according to Theorem 2.1 there exists a solution $u(x)$ which vanishes at $x=\alpha$ but is positive on $(0, \alpha)$. It follows from Lemma 2.2 that $w(x)\neq 0$ on $(0, \alpha)$, for otherwise $u(x)$ would vanish on $(0, \alpha)$. Hence $w(\alpha)=0$, and we have shown that the first conjugate point of $x=0$, if it exists, is the first positive zero of $w(x)$.

Suppose finally that c is a first zero of $w(x)$ on $(0, d)$. Let $z(x)$ be a solution such that $z(x)\not\equiv 0$ and $z(a)=0$ where $c<a<d$. If a differs from c sufficiently little, $z(x)$ will be independent of $w(x)$ and by virtue of Lemma 2.2 must vanish on $(0, c)$. It follows from the definition of the first conjugate point of $x=0$ that $x=0$ possesses a first conjugate point $x=\alpha$. According to the result of the preceding paragraph $c=\alpha$, and we have shown that the first positive zero of $w(x)$, if it exists, is the first conjugate point of $x=0$.

The proof of the theorem is complete.

A solution $w(x)\not\equiv 0$ of the E.E. *such that for each solution $y(x)$ independent of $w(x)$,*

$$\lim_{x=0} \frac{w(x)}{y(x)} = 0$$

will be called a focal solution belonging to $x=0$.

As we have seen when $x=0$ is not its own first conjugate point, the first

* See Bôcher [1].

positive zero of a focal solution is identical with the first conjugate point of $x=0$.

We can now establish the analogue of the Jacobi necessary condition.

THEOREM 2.3. *If* $[0, b]$ *affords a minimum limit to* J *among* A*-admissible curves, there can be no point conjugate to* $x=0$ *on the interval* $[0, b)$.

Suppose $x=\alpha$ on $[0, b)$ is the first conjugate point of $x=0$. By virtue of the definition of such a conjugate point there will exist a solution $v(x)\not\equiv 0$ of the E.E. which vanishes at least twice on $(0, b)$, say at $x=x_0$ and $x=x_1$, where

$$0 < x_0 < x_1 < b.$$

But $[x_0, b]$ affords a minimum to J among neighboring curves of class D^1 which join its end points. Hence a point $x=x_1$ conjugate to $x=x_0$ cannot exist on (x_0, b).

The theorem follows at once.

3. **The Hilbert integral.**† Suppose there is no point conjugate to $x=0$ on $[0, b)$. The solution $w(x)$ such that

$$w(b) = 0, \qquad w'(b) = -1,$$

will not vanish on $(0, b)$. The family of extremals

$$y = \alpha w(x) \qquad (\alpha \text{ constant}) \tag{3.1}$$

will form a field in the region S of the xy-plane for which $0<x<b$. The slope $\lambda(x, y)$ of the extremal of the field through the point (x, y) is given by the equation

$$\lambda(x, y) = y\frac{w'(x)}{w(x)} \qquad (0 < x < b). \tag{3.2}$$

The *Hilbert integral* corresponding to the field (3.1) is a line integral of the form

$$H = \int [f(x, y, \lambda) - \lambda f_{y'}(x, y, \lambda)]dx + f_{y'}(x, y, \lambda)dy, \tag{3.3}$$

where $\lambda=\lambda(x, y)$. The equation (3.1) determines a transformation from the variables x, y to variables x, α. For x, y on S this transformation is one-to-one. In terms of x, α the Hilbert integral takes the form

$$H^* = \int A(x, \alpha)dx + B(x, \alpha)d\alpha, \tag{3.4}$$

† For a treatment of the Hilbert integral for non-singular functionals see Morse [1], p. 13 ff.

where

$$A(x, \alpha) = f[x, \alpha w(x), \alpha w'(x)],$$
$$B(x, \alpha) = w(x) f_{y'}[x, \alpha w(x), \alpha w'(x)].$$

When f is given by (1.1), the line integral (3.4) reduces to

$$H^* = \int \alpha^2(rw'^2 + 2qww' + pw^2)dx + 2\alpha(rww' + qw^2)d\alpha. \tag{3.5}$$

Let S^* denote the region of the $x\alpha$-plane in which $0 < x < b$. Equation (3.1) defines a one-to-one transformation of S^* into S. Let S_b^* be the point set sum of S^* and the line $x = b$ in the $x\alpha$-plane. Under the transformation (3.1) every point (x, α) of the line $x = b$ is carried into the point $(b, 0)$ of the xy-plane.

The functions $A(x, \alpha)$ and $B(x, \alpha)$ are continuous in x and α for (x, α) on S_b^* and, as is well known (cf. Morse [1], p. 13 ff), there exists a function $K(x, \alpha)$ of class C^1 in x, α on S_b^* such that

$$dK = Adx + Bd\alpha.$$

Let γ be a continuous curve on S_b^* in the $x\alpha$-plane of the form

$$x = x(t), \qquad \alpha = \alpha(t) \qquad (t_1 \leqq t \leqq t_2).$$

If γ is rectifiable, the line integral (3.5) along γ exists as a Lebesgue line integral.† If further $x(t)$ and $\alpha(t)$ are absolutely continuous, this line integral equals the Lebesgue integral

$$\int_{t_1}^{t_2} \left\{ A[x(t), \alpha(t)] \frac{dx}{dt} + B[x(t), \alpha(t)] \frac{d\alpha}{dt} \right\} dt.$$

Moreover H^* is independent of paths γ joining two fixed points in S_b^* provided $x(t)$ and $\alpha(t)$ are absolutely continuous. For

$$\frac{dK}{dt} = A\frac{dx}{dt} + B\frac{d\alpha}{dt}$$

almost everywhere on γ, and since $K[x(t), \alpha(t)]$ is absolutely continuous in t,

$$K[x(t), \alpha(t)]_{t_1}^{t_2} = \int_{t_1}^{t_2} \left[A\frac{dx}{dt} + B\frac{d\alpha}{dt} \right] dt = \int_\gamma Adx + Bd\alpha.$$

Thus along a curve γ for which $x(t)$ and $\alpha(t)$ are absolutely continuous and $0 < x \leqq b$, the line integral H^* given by (3.5) depends only on the end points of γ.

† de la Vallée Poussin [1], p. 383.

Referring to (3.5) we see that for b fixed

$$B(b, \alpha) \equiv 0.$$

It follows that

$$K(b, \alpha) \equiv K(b, 0). \tag{3.6}$$

Thus for curves γ which start as a fixed point P on S^* and terminate at an arbitrary point Q on the line $x=b$, the value of H^* is independent of Q.

When interpreted in terms of the integral H in the xy-plane, the foregoing theory yields the following.

Let S_b be the point set sum of S and the point $(b, 0)$ in the xy-plane. Under the transformation (3.1) S_b^* goes into S_b. Let g be a curve on S_b of the form

$$x = x(t), \qquad y = y(t) \qquad (t_1 \leqq t \leqq t_2),$$

where $x(t)$ and $y(t)$ are absolutely continuous. If g lies on S, the Hilbert integral exists and depends only on the end points of g. If g terminates at the point $(b, 0)$, H still exists and depends only on its first end point provided

$$\frac{y(t)}{w[x(t)]} = \alpha(t)$$

is bounded as t tends to t_2. This follows from (3.6) and the continuity of $K(x, \alpha)$.

An extension of the Weierstrass formula. Let $y(x)$ be an A-admissible curve and $u(x)$ a solution of the E.E. not null neighboring $x=0$. Set

$$S[y, u] = -\frac{y^2}{u}(ru' + qu). \tag{3.7}$$

If, in particular, $u(x)$ is a solution which vanishes at $x=b$ but is not identically null, we set

$$S[y, u] = s[y, b],$$

observing that $S[y, u]$ does not depend upon the particular choice of a non-null solution $u(x)$ which vanishes at $x=b$. We term $s[y, b]$ the *singularity function* belonging to $x=b$.

Upon recalling the definition of the Weierstrass E-function† one finds that for our functional J,

$$E[x, y, \lambda, \mu] = r(x)(\lambda - \mu)^2.$$

We are ready for the following extension of the Weierstrass formula.

† Morse [1], p. 5.

THEOREM 3.1. *If $y(x)$ is A-admissible on $[0, b]$ and of class C^1 neighboring $x=b$, if there is no point conjugate to $x=0$ on $[0, b)$, then*

$$(3.8)\quad J(y)\Big|_e^b = \int_e^b E[x, y(x), y'(x), p(x, y(x))]dx + s[y(e), b], \qquad (0 < e < b),$$

where $p(x, y)$ is the slope at the point (x, y) of the field F of extremals through the point $x=b$.

We set up the Hilbert integral H for the field F. Let g denote the curve $y=y(x)$ of the theorem. By virtue of the fact that g is of class C^1 neighboring $x=b$, H exists as a line integral when taken along g from $x=e$ to $x=b$. Let the value of H when taken along g from $x=e$ to $x=b$ be denoted by $H_y\big|_e^b$. Recalling the definition of the Weierstrass E-function, one then sees that

$$J(y)\Big|_e^b = \int_e^b E dx + H_y\Big|_e^b,$$

where the arguments are those of (3.8). We shall prove that

$$(3.9)\qquad H_y\Big|_e^b = s[y(e), b].$$

Let δ denote the line $x=e$. To establish (3.8) we consider the closed curve obtained by following δ from $y=0$ to $y=y(e)$, the curve $y=y(x)$ from $x=e$ to $x=b$, and finally the x-axis from $x=b$ to $x=e$. Taking the Hilbert integral H along this closed curve we find that

$$(3.10)\qquad H_\delta\Big|_{y=0}^{y=y(e)} + H_y\Big|_{x=e}^{x=b} = 0.$$

But upon turning to the explicit definition of H we see that

$$(3.11)\qquad H_\delta\Big|_0^{y(e)} = \int_0^{y(e)} 2\frac{rw' + qw}{w}\, y\, dy = -s[y(e), b].$$

Equation (3.9) is a consequence of (3.10) and (3.11), and (3.8) is proved.

4. The singularity condition. The following example will show the inadequacy of the classical conditions.

EXAMPLE 4.1. Consider the functional

$$J = \int_e^1 \left(\frac{y'^2}{x} - \frac{y^2}{x^3}\right) dx.$$

The Euler equation is

$$\frac{d}{dx}\left(\frac{y'}{x}\right)+\frac{y}{x^3}=0$$

and possesses the linearly independent solutions x and $x\log x$. The solution x is a "focal solution" belonging to $x=0$, so that $x=0$ has no first conjugate point in the sense of §2. Except for the singularity at $x=0$, the analogues of the ordinary conditions that the segment $[0, 1]$ afford a minimum limit to J are satisfied. Nevertheless an easy computation shows that along the curve $y=x(x-1)$,

$$\lim_{\epsilon=0} J(y)\Big|_{\epsilon}^{1}=-\tfrac{1}{2}.$$

The minimum limit thus fails to exist along $[0, 1]$. We are accordingly led to seek further necessary and sufficient conditions for a minimum.

The following theorem introduces a new and important necessary condition.

THEOREM 4.1. *In order that* $[0, b]$ *furnish a minimum limit to* J, *it is necessary that*

$$\liminf_{x=0} s[y(x), b]\geqq 0 \tag{4.0}$$

for each A*-admissible curve for which*

$$\liminf_{\epsilon=0} J(y)\Big|_{\epsilon}^{b} \tag{4.1}$$

is finite.

We shall first prove the theorem for the case in which $y(x)$ is of class C^1 neighboring $x=b$, making this assumption in order that Theorem 3.1 may be applied. Theorem 4.1 will then follow for an arbitrary A-admissible curve g joining the end points of $[0, b]$, since g can be replaced by an A-admissible curve of class C^1 neighboring $x=b$ without altering g except in the neighborhood of $x=b$.

We suppose that the limit (4.1) is finite for the given curve $y=y(x)$. There accordingly exists a decreasing sequence of constants e_n such that

$$\lim_{n=\infty} e_n=0 \qquad (0<e_n<b)$$

and

$$\liminf_{\epsilon=0} J(y)\Big|_{\epsilon}^{b}=\lim_{n=\infty} J(y)\Big|_{e_n}^{b}.$$

We set $y(e_n)=c_n$ and let g_n denote the curve obtained by following an extremal from the point $(b, 0)$ to the point (e_n, c_n) and the curve $y=y(x)$ from (e_n, c_n) to $(0, 0)$. We apply (3.8) to g_n indicating the evaluation of J along g_n by adding the subscript n. We see that

$$J_n\Big|_{e_n}^{b} = s[c_n, b]. \tag{4.2}$$

But g_n is A-admissible so that

$$J_n\Big|_{e_n}^{b} + \lim_{r=\infty} J(y)\Big|_{e_n+r}^{e_n} \geqq 0 \qquad (r>0).$$

Upon letting n become infinite we infer that

$$\lim_{n=\infty} J_n\Big|_{e_n}^{b} \geqq 0.$$

Combining this result with (4.2) we see that

$$\lim_{n=\infty} s[c_n, b] \geqq 0. \tag{4.2$'$}$$

We shall conclude the proof of the theorem by showing that

$$\lim_{n=\infty} s[c_n, b] = \liminf_{x=0} s[y(x), b]. \tag{4.3}$$

To that end we apply (3.8) to the curve $y=y(x)$ setting $e=e_n$. We find that

$$\lim_{n=\infty} \int_{e_n}^{b} E dx \tag{4.4}$$

exists and is finite. From the fact that $E\geqq 0$ we infer that

$$\lim_{e=0} \int_{e}^{b} E dx \tag{4.5}$$

exists, is finite, and equals the limit (4.4). Because of the finiteness of these limits, (3.8) yields the relations

$$\liminf_{e=0} J(y)\Big|_{e}^{b} = \lim_{e=0} \int_{e}^{b} E dx + \liminf_{e=0} s[y(e), b],$$

$$\lim_{n=\infty} J(y)\Big|_{e_n}^{b} = \lim_{e=0} \int_{e}^{b} E dx + \lim_{n=\infty} s[c_n, b],$$

from which (4.3) follows as stated.

Condition (4.0) is a consequence of (4.2) and (4.3), and the proof is complete.

EXAMPLE 4.2. The condition (4.1) is independent of the analogues of the classical Jacobi and Weierstrass conditions. For in the example at the beginning of this section the point $x=0$ has no conjugate point, and the Weierstrass condition holds on $(0, b]$ since $r(x)>0$ on $(0, b]$.

Let g denote the curve $y=x(x-1)$. The curve g is A-admissible on $[0, 1]$. Moreover the limit (4.1) is finite along g. Observe that $x \log x$ is a solution $u(x)$ of the E.E. which vanishes when $x=1$, and that when $y=x(x-1)$,

$$s[y, 1] = -r\frac{u'}{u}y^2 = -\frac{(1+\log x)}{\log x}(x-1)^2.$$

We see that

$$\lim_{x=0} s[y(x), 1] = -1,$$

so that the condition is not satisfied.

EXAMPLE 4.3. The provision in the theorem that the limit (4.1) be finite cannot be removed as an examination of the functional

$$\int_e^\pi (y'^2 - y^2)dx$$

will show.

It is clear that the limit (4.1) is not finite for the curve $y=x^{1/2}(x-\pi)$. The singularity function belonging to $x=\pi$ takes the form

$$s[y, \pi] = -y^2 \cot x.$$

Moreover $y=x^{1/2}(x-\pi)$ is an A-admissible curve g for which

$$\lim_{x=0} s[y(x), \pi] = -\pi^2,$$

so that (4.0) is not satisfied. Nevertheless we shall see in the next section that $[0, \pi]$ affords a minimum limit for this functional.

The condition of Theorem 4.1 that (4.0) hold for each A-admissible curve for which the limit (4.1) is not $+\infty$ will be termed the *singularity condition* belonging to the segment $[0, b]$.

5. **Sufficient conditions.** Let a and b be two constants such that $0<a<b<d$, and let g be an arbitrary extremal $y=g(x)$ defined on the interval $a \leqq x \leqq b$. Let $y=z(x)$ be an absolutely continuous curve which joins the end points of g and on which z'^2 is in L.

If there is no point conjugate to $x=a$ on $[a, b]$,

$$J(z)\Big|_a^b \geqq J(g)\Big|_a^b. \tag{5.1}$$

The proof of (5.1) follows classical lines. One covers g with a field F of extremals without singularity neighboring g, and makes use of the Hilbert integral theory as extended in §3. The proof of (5.1) is then completed in the usual way.

We turn to the question of the equality in (5.1). Let $p(x, y)$ be the slope of F at the point (x, y). The classical formula of Weierstrass

$$\Delta J = \int_a^b E dx$$

holds and shows that the equality in (5.1) prevails only if

$$\frac{dz}{dx} = p[x, z(x)] \tag{5.2}$$

almost everywhere on $[a, b]$. But $z(x)$ is absolutely continuous so that it follows from (5.2) that

$$z(x) - z(a) = \int_a^x p[x, z(x)]dx.$$

We infer that $z(x)$ is of class C^1. Returning to (5.2) we see that $z = z(x)$ must then be an extremal of F and hence identical with g.

We come to a principal theorem.

THEOREM 5.1. *Necessary and sufficient conditions that* $[0, b]$ *afford a minimum limit to* J *are that* $[0, b)$ *contain no point conjugate to* $x = 0$, *and that the singularity condition belonging to* $[0, b]$ *be satisfied.*

That the conditions of the theorem are necessary has already been established.

To prove the conditions of the theorem sufficient we shall begin with an A-admissible curve $y = y(x)$ of class C^1 neighboring $x = b$ and prove that

$$\liminf_{\epsilon=0} J(y)\Big|_a^b \geqq 0. \tag{5.3}$$

The relation (5.3) will then follow for an A-admissible curve h not of class C^1 neighboring $x = b$. For one can replace a segment of h terminating at $x = b$ on which x varies sufficiently little by an extremal joining its end points, obtaining thereby a new curve k of class C^1. Moreover it follows from (5.1) that

$$J(h)\Big|_a^b \geqq J(k)\Big|_a^b,$$

since there is no point conjugate to $x = b$ on the interval on which h is altered.

But in case $y(x)$ is of class C^1 neighboring $x=b$ we can use formula (3.8) and infer that (5.3) holds.

The proof of the theorem is complete.

The following theorem is of importance in the applications.

THEOREM 5.2. *If for every value of c for which $0<c<b$, and for every curve k which is A-admissible on $[0, b]$,*

$$\liminf_{\epsilon=0} J(k)\Big|_\epsilon^c \geqq 0,$$

then $[0, b]$ will afford a minimum limit to J along curves which are A-admissible on $[0, b]$.

Let A and B be the end points of $[0, b]$ and let g be an A-admissible curve which joins A to B. Let P be a point on g near B but not B. Let h be an extremal which joins P to B. If P is sufficiently near B, h will afford a minimum to J among A-admissible curves which join its end points. Let γ be a curve which follows g from A to P and h from P to B. Holding P fast let γ be altered by replacing B by a point $(c, 0)$ on the x-axis near B with $0<c<b$ and replacing h by an extremal joining P to $(c, 0)$. For the resulting curve μ,

$$\liminf_{\epsilon=0} J(\mu)\Big|_\epsilon^c \geqq 0 \tag{5.4}$$

by hypothesis. If the limit (5.4) is finite, we see that

$$\lim_{c=b}\ \liminf_{\epsilon=0} J(\mu)\Big|_\epsilon^c = \liminf_{\epsilon=0} J(\gamma)\Big|_\epsilon^b \geqq 0.$$

If the limit (5.4) is $+\infty$, we see that

$$\liminf_{\epsilon=0} J(\gamma)\Big|_\epsilon^b = +\infty.$$

The theorem follows readily.

In the following theorem and proof each inferior limit is an inferior limit as x tends to zero.

THEOREM 5.3. *If $[0, e]$ affords a minimum limit to J for any sufficiently small positive constant e, and if there exists a solution $u(x)$ of the* E.E. *for which the* lim inf $u(x)\neq 0$, *J assumes a minimum limit on any interval $[0, b]$ such that $[0, b)$ contains no point conjugate to $x=0$.*

Let $w(x)$ be a focal solution. We begin the proof by establishing statements (a) and (b).

(a) *If $v(x)$ is a solution of the* E.E. *independent of $w(x)$,* $\liminf v(x) \neq 0$.

If $\liminf w(x) = 0$, (a) is true, for otherwise $\liminf u(x)$ would be null. If $\liminf w(x) \neq 0$, $\liminf v(x)$ must be infinite since $\lim w/v = 0$. Thus (a) is true in all cases.

(b) *If $y(x)$ is A-admissible,* $\liminf J(y)\Big|_x^b$ *is finite, and $v(x)$ is independent of $w(x)$, then*

$$\liminf S[y(x), v(x)] \geqq 0. \tag{5.5}$$

If the constant e is sufficiently small, the solution $z(x)$ for which

$$z(e) = 0, \qquad z'(e) = -1$$

will have a non-null inferior limit. Since $[0, e]$ affords a minimum limit to J,

$$\liminf S[y(x), z(x)] \geqq 0 \tag{5.6}$$

by virtue of Theorem 4.1. But

$$S[y, v] - S[y, z] = \frac{y^2 r}{zv}\,[z'v - v'z] = \frac{ky^2}{zv}, \tag{5.7}$$

where k is a constant. Now $\liminf v(x) \neq 0$ by virtue of (a), and it follows from (5.7) that

$$\liminf S[y, v] = \liminf S[y, z]. \tag{5.8}$$

Statement (b) follows from (5.6) and (5.8).

Theorem 5.1 now shows that $[0, b]$ affords a minimum limit to J provided $x = 0$ is not conjugate to $x = b$. That the theorem is valid in case $x = 0$ is conjugate to $x = b$ follows from Theorem 5.2, and the proof is complete.

Example 4.2 shows that statement (b) is false if the phrase "$v(x)$ is independent of $w(x)$" is deleted.

We introduce the following theorem.

THEOREM 5.4. *If $p \geqq 0$ neighboring $x = 0$, if $q \equiv 0$, and if there is no point on $[0, b)$ conjugate to $x = 0$, J assumes a minimum limit on $[0, b]$ among A-admissible curves.*

Let e be a positive constant such that $p \geqq 0$ on $(0, e)$. Let $u(x)$ be a solution of the E.E. such that $u(e) = 0$ and $u(x) > 0$ on $(0, e)$. Observe that

$$J(u)\Big|_x^e = -r(x)u(x)u'(x) > 0.$$

Hence $u'(x) < 0$ on $(0, e)$ and $\liminf u(x) \neq 0$. We infer that (5.5) holds as stated. The theorem then follows from Theorem 5.3.

We conclude with the following theorem.

THEOREM 5.5. *The minimum limits affirmed to exist in Theorems* 5.3 *and* 5.4 *are proper unless* $x=b$ *is a conjugate point of* $x=0$. *If* $x=b$ *is a conjugate point of* $x=0$, *the minimum limits may be proper or improper.*

Suppose $x=b$ is not conjugate to $x=0$. Under the hypotheses of the preceding theorems if $b<c<d$, and c differs sufficiently little from b, the functional J will assume a minimum limit on $[0, c]$ among curves which are A-admissible on $[0, c]$. A curve $y=u(x)$ which is A-admissible on $[0, b]$ and which gives a null inferior limit to J must be an extremal. A curve g which follows $y=u(x)$ from $x=0$ to $x=b$ and the x-axis from $x=b$ to $x=c$ will be A-admissible on $[0, c]$ and give a null inferior limit to J. The curve g must then be an extremal. This is possible only if $u(x)\equiv 0$. Hence $[0, b]$ affords a proper minimum limit to J unless $x=b$ is conjugate to $x=0$, and the proof is complete.

The segment $[0, \pi]$ of the x-axis furnishes an improper minimum limit to the integral

$$\int_0^\pi (y'^2 - y^2)dx.$$

On the other hand we shall see in §8 that there are examples in which $x=b$ is conjugate to $x=0$ and $[0, b]$ furnishes a proper minimum limit to J.

II. ONE SINGULAR END POINT; REGULAR CASE

6. **The integral.** Part II will be devoted to the study of integrals of the form

$$\int_a^b [x^\alpha g(x)y'^2 - x^{\alpha-2}h(x)y^2]dx. \tag{6.1}$$

Here α is any real number, $g(x)$ and $h(x)$ are real analytic functions of x on the interval $[0, d)$, and $g>0$ on this interval. The corresponding E.E. may be given the form

$$x^2gy'' + x(\alpha g + xg')y' + hy = 0. \tag{6.2}$$

It will be observed that $x=0$ is a regular* singular point of this differential equation.

Conversely, any differential equation of the form

$$x^2\lambda(x)y'' + x\mu(x)y' + \nu(x)y = 0 \tag{6.3}$$

* Most standard works on ordinary differential equations give a treatment of the regular singular point which is based on the methods of J. Frobenius. G. D. Birkhoff [1] gives an alternative simplified treatment of this question.

in which λ, μ, ν are real and analytic on $[0, d)$, and λ is positive on $[0, d)$, after multiplication by the factor

$$x^{\alpha-2}\frac{g(x)}{\lambda(x)}, \tag{6.4}$$

where

$$\alpha = \frac{\mu(0)}{\lambda(0)}, \qquad g(x) = \exp\left[\int_0^x \frac{1}{x}\left(\frac{\mu(x)}{\lambda(x)} - \frac{\mu(0)}{\lambda(0)}\right)dx\right], \tag{6.5}$$

becomes the E.E. of an integral of the form (6.1) in which α and g are defined by (6.5) and

$$h(x) = \frac{g(x)\nu(x)}{\lambda(x)}.$$

With the equation (6.2) one associates the so-called indicial equation

$$\rho^2 - \rho(1-\alpha) + \beta = 0, \tag{6.6}$$

where

$$\beta = \frac{h(0)}{g(0)}.$$

The roots σ and τ of (6.6) are called *indicial roots* associated with the point $x=0$. They satisfy the relations

$$\sigma + \tau = 1 - \alpha, \qquad \sigma\tau = \beta. \tag{6.7}$$

When σ and τ are real, we shall suppose that

$$\sigma \geqq \tau.$$

As is well known,* there exist two linearly independent solutions $w(x)$ and $v(x)$ of (6.2) of the respective forms

$$w(x) = x^{\sigma}a(x), \qquad v(x) = x^{\tau}b(x) + cw(x)\log x, \qquad (0 < x < d), \tag{6.8}$$

where $a(x)$ and $b(x)$ are analytic in x on $[0, d)$,

$$a(0)\cdot b(0) \neq 0,$$

* See, for example, Bieberbach [1], p. 214.

and c is a constant. If $\sigma-\tau$ is not an integer, $c=0$, and if $\sigma=\tau$, $c\neq 0$. If $\sigma-\tau$ is an integer $\neq 0$, c may or may not be zero depending on the differential equation.

Two cases arise according as the indicial roots are real or imaginary. Suppose the roots are of the form $k\pm hi$ where k and h are real and $h\neq 0$. The difference

$$\sigma-\tau=\pm 2hi$$

is not an integer so that $c=0$. The solutions (6.8) may then be replaced by the linearly independent solutions

$$\begin{aligned} W(x) &= x^k[A(x)\cos\log x^h + B(x)\sin\log x^h], \\ V(x) &= x^k[B(x)\cos\log x^h - A(x)\sin\log x^h], \quad (0<x<d), \end{aligned} \tag{6.9}$$

where $A(x)$ and $B(x)$ are real and analytic on $[0, d)$ with

$$A^2(0)+B^2(0)\neq 0. \tag{6.10}$$

7. **Focal solutions and the singularity function.** We shall prove the following theorem.

THEOREM 7.1. *A necessary and sufficient condition that $x=0$ be its own first conjugate point is that the indicial roots be imaginary.*

If the roots are imaginary, a solution of the E.E. is given by $W(x)$ in (6.9). Since $W(x)$ vanishes infinitely often near $x=0$, $x=0$ must be its own first conjugate point. On the other hand, if the exponents of σ and τ are real, $w(x)$ in (6.8) gives a solution which vanishes near $x=0$ for no positive value of x so that in this case $x=0$ cannot be its own first conjugate point.

The proof of the theorem is complete.

We continue with the following theorem.

THEOREM 7.2. *If the indicial roots are real and $\sigma\geqq\tau$, the solution $w=x^\sigma a(x)$ of* (6.8) *is a focal solution of the* E.E. *and the conjugate points of $x=0$ are the zeros of $a(x)$.*

In terms of the solutions $w(x)$ and $v(x)$ of (6.8) we see that

$$\lim_{x=0}\frac{w(x)}{v(x)}=0.$$

If then $y(x)$ is any solution independent of $w(x)$,

$$\lim_{x=0}\frac{w(x)}{y(x)}=0,$$

so that $w(x)$ is a focal solution of the E.E. That the conjugate points of $x=0$ are the zeros of $a(x)$ follows from Theorem 2.2.

We shall next verify three formulas useful in the evaluation of the singularity function $s[y, b]$ defined in §4. These formulas concern the solutions $w(x)$ and $v(x)$ of (6.8) together with any solution $u(x)$ of the form

$$u = c_1 w + c_2 v \qquad (c_2 \neq 0).$$

These formulas are as follows:

$$\frac{w'}{w} = \frac{1}{x}[\sigma + o(1)], \tag{7.1}$$

$$\frac{v'}{v} = \frac{1}{x}[\tau + o(1)], \tag{7.2}$$

$$\frac{u'}{u} = \frac{1}{x}[\tau + o(1)]. \tag{7.3}$$

Formula (7.1) is verified at once.

To establish (7.2) observe that

$$\frac{v'}{v} = \frac{1}{x}\left[\tau + \frac{xb'(x) + ck(x) + cxk'(x)\log x}{b(x) + ck(x)\log x}\right], \tag{7.4}$$

where

$$k(x) = x^{\sigma-\tau}a(x).$$

Upon separating the cases $ck(0)=0$ and $ck(0)\neq 0$, (7.2) appears as a consequence of (7.4).

Finally we see that

$$\frac{u'}{u} = \frac{c_1 w' + c_2 v'}{c_1 w + c_2 v} = \frac{c_1 \dfrac{w'}{w}\dfrac{w}{v} + c_2 \dfrac{v'}{v}}{c_1 \dfrac{w}{v} + c_2}. \tag{7.5}$$

Formula (7.3) follows from (7.5) upon making use of (7.1) and (7.2) together with the fact that w/v tends to zero as x tends to zero.

8. **Sufficient conditions in the regular singular case.** We consider the functional

$$J = \int [g(x)x^{\alpha}y'^2 - h(x)x^{\alpha-2}y^2]dx \tag{8.0}$$

of §6 excluding the trivial case $h(x)\equiv 0$. We are concerned with conditions that $[0, b]$ afford a minimum limit to J.

The exponent m. It will be convenient to rewrite J in the form

$$J = \int [g(x)x^{\alpha}y'^2 - k(x)x^{m-2}y^2]dx, \tag{8.1}$$

where $k(x)$ is analytic on the interval $[0, d)$ and $k(0) \neq 0$. We see that $m \geqq \alpha$.

Besides the conjugate point condition our sufficient conditions will involve the sign of $k(0)$ and the exponent m. To distinguish between proper and improper minima we shall need the roots σ and τ. Recall that σ and τ are the roots of the equation

$$\rho^2 - \rho(1 - \alpha) + \beta = 0 \tag{8.2}$$

with $\sigma \geqq \tau$ by convention when the roots are real.

We shall make use of the following lemma.

LEMMA 8.1. *If $y(x)$ is A-admissible and $x^{\alpha}y'^2(x)$ is in L on $[0, b]$, $x^{\alpha-1}y^2(x)$ tends to zero as a limit as x tends to zero.*

The lemma is trivial if $\alpha \geqq 1$. Suppose then that $\alpha < 1$. Let e and x be so chosen that $0 < e < x < b$. We note that

$$\left[y(x)\Big|_e^x\right]^2 = \left[\int_e^x y'dx\right]^2.$$

Upon using the Schwartz inequality we find that

$$\left[y(x)\Big|_e^x\right]^2 = \left[\int_e^x x^{-\alpha/2}(x^{\alpha/2}y')dx\right]^2 \leqq \int_e^x \frac{dx}{x^{\alpha}} \int_e^x x^{\alpha}y'^2dx.$$

If we let e tend to zero as a limit, this relation leads to the inequality

$$(1 - \alpha)x^{\alpha-1}y^2(x) \leqq \int_0^x x^{\alpha}y'^2(x)dx,$$

from which the lemma follows as stated.

We continue with the following lemma.

LEMMA 8.2. *If $y(x)$ is A-admissible and $x^{\alpha}y'^2(x)$ is in L on $[0, b]$, and if the indicial roots are real,*

$$\lim_{x=0} S[y(x), u(x)] = 0$$

for every A-admissible curve $y = y(x)$ and solution $u(x) \not\equiv 0$.

Upon referring to the definition of the singularity function and making use of (7.1), (7.2), and (7.3) we see that

$$|S[y(x), u(x)]| \leqq x^{\alpha-1}y^2(x)O(1).$$

Lemma 8.2 follows from Lemma 8.1.

We come to a principal theorem.

THEOREM 8.1. *Necessary and sufficient conditions that* $[0, b]$ *afford a minimum limit to* J *among* A*-admissible curves are that* $[0, b)$ *contain no point conjugate to* $x=0$ *and that* $m>1$ *whenever* $k(0)>0$.

We have already proved that it is necessary that $[0, b)$ contain no point conjugate to $x=0$.

That the conditions of the theorem are sufficient when $k(0)<0$ follows from Theorems 5.4 and 2.3.

The case $m \leqq 1$, $k(0)>0$. We shall show that $[0, b)$ fails to afford a minimum limit to J in this case. Observe that the term in J of the form

$$-\int_x^b k(x)x^{m-2}y^2dx \tag{8.3}$$

will decrease without limit on a curve of the form $y=$ constant $\neq 0$, as x tends to zero.

An A-admissible curve g for which

$$\lim_{e=0} J(g)\Big|_e^b \leqq -1 \tag{8.4}$$

can accordingly be constructed from an infinite succession of straight line segments

$$g_0, g_1, g_2, \cdots$$

as follows. We suppose g_0 terminates at the point $(b, 0)$, has the slope -1 and a length equal to b. Let (a_n, b_n) denote the initial point of g_n and suppose that $0<a_n<a_{n-1}$. For n odd, g_n shall terminate at (a_{n-1}, b_{n-1}) and have the slope zero. Its initial point (a_n, b_n) shall be so chosen that

$$J(g)\Big|_{a_n}^b < -1.$$

For n odd, g_{n+1} shall terminate at (a_n, b_n), have a slope b_n/a_n and a length equal to half the distance of (a_n, b_n) from the origin.

It is clear that g exists, is A-admissible and that on g (8.4) holds. If $[0, b]$ affords a minimum limit to J among A-admissible curves, it is accordingly necessary that $m>1$.

The case $m>1$, $k(0)>0$. In this case the integral (8.4) has a finite limit as e tends to zero. We distinguish between the case in which the limit

$$\lim_{e=0} \int_e^b x^\alpha y'^2 dx \tag{8.5}$$

is finite and the case in which it is infinite. In the latter case it is clear that

$$\liminf_{\epsilon=0} J(y)\Big|_{\epsilon}^{b} > 0, \tag{8.6}$$

as desired. In case the limit in (8.5) is finite and $[0, b)$ contains no point conjugate to $x=0$, it follows from Lemma 8.2 that

$$\lim_{x=0} s[y(x), b] = 0.$$

We see from Theorem 3.1 that (8.6) holds in this case as well.

The proof of the theorem is complete.

We introduce the following lemma.

LEMMA 8.3. *If* $[0, b]$ *affords a minimum limit to* J *among* A*-admissible curves, the smaller exponent* τ *is never positive.*

We distinguish three cases.

CASE I. $m>\alpha$. In this case $\beta=0$ in (8.2). But $\beta=\sigma\tau$, and the lemma is clearly true.

CASE II. $m=\alpha$, $k(0)<0$. Here $\beta<0$ and the lemma is again true.

CASE III. $m=\alpha$, $k(0)>0$. By virtue of the theorem $m>1$ in this case. Hence

$$\alpha = 1 - \sigma - \tau > 1, \qquad \beta = \sigma\tau > 0,$$

and we infer that $\tau<0$.

The proof of the lemma is complete.

The following theorem completes the preceding theory.

THEOREM 8.2. *Necessary and sufficient conditions that* $[0, b]$ *afford an improper minimum limit to* J *among* A*-admissible curves are that* $x=b$ *be the first conjugate point of* $x=0$, *and that* $m>1>\alpha$ *in case* $k(0)>0$, *and that* $\alpha=m$ *or* $\alpha<1$ *in case* $k(0)<0$.

That it is necessary that $x=b$ be the first conjugate point of $x=0$ is proved as under Theorem 5.5. We assume the facts affirmed under Theorem 8.1 and continue with a proof of the following.

(a) *If* $x=b$ *is the first conjugate point of* $x=0$, *if* $k(0)>0$ *and* $m>1$, *the minimum limit assumed by* J *on* $[0, b]$ *is improper if and only if* $\alpha<1$.

We shall first prove that an improper minimum limit in (a) implies that $\alpha<1$. To that end observe that such a limit implies that the focal solution

$$w(x) = x^{\sigma}a(x)$$

is A-admissible, and hence $\sigma>0$. To continue we distinguish between two cases.

CASE I. $m=\alpha$. In this case $\sigma\tau>0$. But $\tau\leqq 0$ by virtue of Lemma 8.3 so that Case I is impossible.

CASE II. $m>\alpha$. In this case the indicial roots are $1-\alpha$ and 0. Thus $\alpha<1$ since $\sigma>0$.

Hence an improper minimum limit in (a) implies that $\alpha<1$.

Conversely, we shall assume that $\alpha<1$ in (a) and prove that $y=w(x)$ is A-admissible and

$$\lim_{x=0} J(w)\Big|_x^b = 0. \tag{8.7}$$

First, observe that

$$1-\sigma-\tau<1$$

when $\alpha<1$, so that $\sigma>0$. It follows from (8.7) that

$$J(w)\Big|_x^b = -g(x)x^{\alpha}ww' = O[x^{\alpha}x^{\sigma}x^{\sigma-1}]$$

$$= O[x^{\sigma-\tau}]. \tag{8.8}$$

But $\tau\leqq 0$ by virtue of Lemma 8.3 so that $\sigma-\tau>0$. Relation (8.7) follows from (8.8), and the proof of (a) is complete.

We continue with a proof of statement (b).

(b) *If $x=b$ is the first conjugate point of $x=0$ and if $k(0)<0$, the minimum limit assumed by J on $[0, b]$ is improper if and only if $m=\alpha$ or $\alpha<1$.*

As in the proof of (a) we see that the minimum limit in (b) is improper only if $\sigma>0$. If $m>\alpha$, the indicial roots are $1-\alpha$ and 0, and if further $\sigma>0$, it is necessary that $\alpha<1$. Hence the minimum limit in (b) is improper only if $m=\alpha$ or $\alpha<1$.

If, conversely, $\alpha=m$ and $k(0)<0$, we infer that $\sigma\tau<0$ and hence $\sigma-\tau>0$. Thus (8.7) follows from (8.8) when $\alpha=m$ in (b). But on the other hand if $m>\alpha$, the indicial roots are $1-\alpha$ and 0, and if in addition $\alpha<1$, $\sigma=1-\alpha>0$ so that (8.8) again follows from (8.7).

The proof of (b) and of the theorem is complete.

For completeness we add the following.

The minimum limit zero in Theorem 8.2 is assumed by J only on solutions dependent on focal solutions.

In the regular singular case $m\geqq\alpha$. We shall prove the following theorem.

THEOREM 8.3. *If $m<\alpha$ and $k(0)>0$ in* (8.1), $[0, b]$ *fails to afford a minimum limit to J among A-admissible curves.*

Let c be a positive constant so large that the indicial roots belonging to the E.E. of the functional

$$J^* = \int_0^b [x^\alpha g y'^2 - c x^{\alpha-2} k y^2] dx$$

are imaginary. Under the hypotheses of the theorem there exists a positive constant $\eta<b$ such that

$$k(x) > 0, \qquad cx^{\alpha-m} < 1$$

on $[0, \eta]$. For such a choice of η,

$$J^*(y)\Big|_e^\eta \geqq J(y)\Big|_e^\eta \qquad (0 < e < \eta) \tag{8.9}$$

for every $y(x)$ which is A-admissible on $[0, \eta]$. But since the indicial roots belonging to J^* are imaginary, an A-admissible function $y(x)$ exists for which the left member of (8.9) has a negative inferior limit as e tends to zero.

The theorem follows from (8.9).

9. **Other types of admissibility.** We continue with the regular singular case with the functional (8.0).

Curves α-admissible. A curve $y=y(x)$ which is A-admissible on $[0, b]$ will be said to be *α-admissible* on $[0, b]$ if the limit

$$\lim_{e=0} \int_e^b x^\alpha y'^2 dx \tag{9.1}$$

is finite. It follows from Lemma 8.2 that

$$\lim_{x=0} S[y(x), u(x)] = 0 \tag{9.2}$$

whenever the indicial roots are real, $y(x)$ is A-admissible, and $u(x)$ is a solution of the E.E. not identically zero.

We continue with the following lemma.

LEMMA 9.1. *If $y(x)$ is α-admissible, $x^{\alpha-2}y^2$ is in L on* $[0, b]$ *except at most when $\alpha=1$.*

To prove this lemma we consider the functional

$$I = \int_e^x [x^\alpha y'^2 - K x^{\alpha-2} y^2] dx \qquad (\alpha \neq 1), \tag{9.3}$$

where K is a positive constant. Since $\alpha \neq 1$, the constant K can be chosen so small that the indicial roots corresponding to I are real. We see from (3.8) and (9.2) that

$$\int_0^b x^\alpha y'^2 dx \geqq \lim_{\epsilon=0} K \int_\epsilon^x x^{\alpha-2} y^2 dx.$$

The lemma follows directly.

EXAMPLE 9.1. The lemma would be false were not the case $\alpha = 1$ excepted as is shown by the following example. Neighboring $x = 0$, we suppose that

$$y(x) = [-\log x]^{-1/2}.$$

We find that

$$\int_\epsilon^x xy'^2 dx = -\frac{1}{4}\int_\epsilon^x (\log x)^{-3} d \log x = \frac{1}{8}[\log x]^{-2}\Big|_\epsilon^x,$$

$$\int_\epsilon^x x^{-1}y^2 dx = -\int_\epsilon^x [\log x]^{-1} d \log x = -[\log(-\log x)]_\epsilon^x.$$

From these relations we see that xy'^2 is in L on $[0, 1]$ while $x^{-1}y^2$ is not in L.

We are considering the functional J given by (8.0). When $y(x)$ is α-admissible, it follows from the preceding lemma that the limit

$$\lim_{\epsilon=0} J(y)\Big|_\epsilon^b$$

is finite except at most when $\alpha = 1$. When $\alpha = 1$, this limit may be finite or infinite.

The principal theorem follows.

THEOREM 9.1. *A necessary and sufficient condition that* $[0, b]$ *afford a minimum limit to* J *among* α*-admissible curves is that* $[0, b)$ *contain no point conjugate to* $x = 0$. *This minimum limit is improper if and only if* $x = b$ *is the first conjugate point of* $x = 0$, *if* $\sigma > 0$ *and* $\sigma > \tau$.

The theorem follows readily upon abstracting the appropriate arguments from the proofs in §8. In particular the reader should refer to (8.8) in proving the concluding affirmation.

Curves F-admissible. If

(1) $y(x)$ is continuous on $(0, b]$ and $y(b) = 0$,

(2) $y(x)$ is absolutely continuous and $y'^2(x)$ is in L on each closed subinterval of $(0, b]$, the curve $y = y(x)$ will be said to be *F-admissible* on $[0, b]$. We are concerned with the functional J of (8.0) and (8.1) and state the following theorem.

THEOREM 9.2. *In order that* $(0, b]$ *afford a minimum limit to* J *among* F*-admissible curves, it is sufficient that* $[0, b)$ *contain no point conjugate to* $x=0$ *and that* $\tau<0$ *or* $\sigma=\tau=0$. *Under these conditions the minimum limit is improper if and only if* $x=b$ *is the first conjugate point of* $x=0$ *and* $\sigma>0$.

In case $\tau<0$ or $\sigma=\tau=0$, the solution $v(x)$ of (6.8) is of one sign neighboring $x=0$. If $v(x)$ is chosen so as to be positive, its derivative $v'(x)$ will be negative neighboring $x=0$. It follows that

$$S[y(x), v(x)] \geqq 0 \qquad (x > 0)$$

for all values of x sufficiently near zero. The first statement of the theorem is thus a consequence of (3.8) when $x=b$ is not conjugate to $x=0$, and follows when $x=b$ is conjugate to $x=0$ by the limiting process used in the proof of Theorem 5.2. The second statement of the theorem is a consequence of earlier remarks and of formula (8.8).

The proof of the theorem is complete.

The "general" case in which $(0, b]$ affords a minimum limit to J is described in the preceding theorem. There are known special cases. To give final necessary and sufficient conditions would require the introduction of a new element, the first focal point of the y-axis. For the cases included in the theorem, this focal point is identical with the first conjugate point of $x=0$. We omit further details.

Curves admissible relative to an extremal g. Let $y=u(x)$ be an extremal g which joins the points $(0, 0)$ and (b, c). We assume that $J(u)\Big|_e^b$ tends to a finite limit as e tends to zero. Relative to the extremal g we admit comparison curves of the form

$$y = u(x) + \eta(x),$$

where $\eta(x)$ is X-admissible on $[0, b]$, X symbolizing A, α, or F.

If for each such curve

$$\liminf_{e=0} J(y)\Big|_e^b \geqq J(u)\Big|_0^b,$$

we shall say that g affords a minimum limit to J among the particular class of curves admitted.

A simple integration by parts shows that

$$J(y)\Big|_e^b = J(u)\Big|_e^b + J(\eta)\Big|_e^b + T[\eta(e), u(e)], \tag{9.4}$$

where

$$T[\eta, u] = 2\eta[ru' + qu].$$

If our knowledge of the functional T is sufficient, the preceding theorems about $J(\eta)$ will lead to theorems about $J(y)$. As an example we shall translate Theorem 9.1.

The solution $u(x)$ will vanish at $x=0$ and $J(u)\Big|_\epsilon^b$ will tend to a finite limit if and only if $u(x)$ is a focal extremal and $\sigma>0$. Under these conditions $x^\alpha u'^2$ will be in L on $[0, b]$, and $x^\alpha(u'+\eta')^2$ will be in L on $[0, b]$ if and only if $x^\alpha \eta'^2$ is in L on $[0, b]$.

THEOREM 9.3. *A necessary and sufficient condition that a focal extremal*

$$y = x^\sigma a(x) \qquad (\sigma > 0; 0 \leqq x \leqq b)$$

afford a minimum limit to J relative to curves $y=u(x)+\eta(x)$ for which $\eta(x)$ is α-admissible on $[0, b]$ is that $[0, b)$ contain no point conjugate to $x=0$. This minimum limit is improper if and only if $x=b$ is the first conjugate point of $x=0$ and $\sigma>\tau$. The value of the minimum limit is $g(b)u(b)u'(b)$.

That the conjugate point condition insures a minimum limit will follow from (9.4) and Theorem 9.1 once we have shown that

$$\lim_{\epsilon=0} T[\eta(\epsilon), u(\epsilon)] = 0. \tag{9.5}$$

Observe that

$$T(\eta, u) = \eta(x)x^\alpha x^{\sigma-1}O(1).$$

But η is α-admissible, and it follows from Lemma 8.1 that

$$\eta = x^{(1-\alpha)/2}o(1).$$

Recalling that $\alpha=1-\sigma-\tau$ we see that

$$T(\eta, u) = x^{(\sigma-\tau)/2}o(1),$$

so that (9.5) holds as stated.

The remainder of the theorem follows from (9.4) and Theorem 9.1.

EXAMPLE 9.2. We shall now derive a theorem of H.L.P. With a slight variation which accords with the present notation this theorem is as follows.

If $\mu>4$, $y(0)=0$, $y(1)=1$, $y(x)$ is absolutely continuous and y' is in L^2, then

$$J(y) = \int_0^1 \left[\mu y'^2 - \frac{y^2}{x^2}\right] dx \geqq \mu\sigma.$$

The only case of equality occurs when $y=x^\sigma$.

The indicial equation is

$$\rho^2 - \rho + \frac{1}{\mu} = 0.$$

If $\mu>4$, the roots are real and distinct with $\sigma>0$. The focal solution x^σ vanishes only when $x=0$ so that there is no point conjugate to $x=0$. The minimum limit afforded by x^σ is accordingly proper. Its value is

$$g(1)u(1)u'(1) = \mu\sigma.$$

III. Two singular end points

10. Necessary conditions. We return to a function f of the form (1.1) and a functional of the form

$$J(y)\Big|_{x_1}^{x_2} = \int_{x_1}^{x_2} f(x, y, y')dx, \tag{10.1}$$

where x_1 and x_2 are constants such that

$$0 < x_1 < x_2 < d. \tag{10.2}$$

Functions $y(x)$ which are A-, α-, or F-admissible on $[0, b]$ have already been defined. A function $y(x)$ will be termed A-, α-, or F-admissible respectively on $[0, d]$ if for each constant b on $(0, d)$, $y(x)$ satisfies the conditions for A-, α-, or F-admissibility on $[0, b]$ excepting at most the condition that $y(b)=0$.

The transformation

$$x^* = d - x \tag{10.3}$$

interchanges the points $x=0$ and $x=d$. Functions which result from curves A-, α-, or F-admissible on the interval $[0, b]$ or $[0, d]$ upon applying the transformation (10.3) will be termed A^*-, α^*-, or F^*-admissible respectively on $[0, b]$ or $[0, d]$.

Let X denote one of the symbols A, α, and F, and Y one of the symbols A^*, α^*, and F^*. A function $y(x)$ will be termed (X, Y)-*admissible* on $[0, d]$ if by an alteration on an arbitrarily small neighborhood of $x=d$, $y(x)$ can be made X-admissible on $[0, d]$, and if by an alteration on an arbitrarily small neighborhood of $x=0$, $y(x)$ can be made Y-admissible on $[0, d]$. A curve $y=y(x)$ will be termed (X, Y)-admissible if $y(x)$ is (X, Y)-admissible. We thus have nine different types of (X, Y)-admissibility.

We shall begin by seeking conditions under which

$$\liminf_{x_1=0,\, x_2=d} \int_{x_1}^{x_2} f(x, y, y')dx \geqq 0 \tag{10.4}$$

for curves which are (A, A^*)-admissible on $[0, d]$. If (10.4) holds, we shall say that $[0, d]$ affords a *minimum limit* to J among curves (A, A^*)-admissible on $[0, d]$.

In (10.2) we defined the first conjugate point of $x=0$ on $[0, d)$. Under the transformation (10.3), J will be carried into a functional J^*. Relative to J the first conjugate point of $x=d$ will be defined as the image under (10.3) of the first conjugate point of $x^*=0$ relative to J^*.

A solution $w(x)\not\equiv 0$ of the E.E. such that for each solution $y(x)$ independent of $w(x)$

$$\lim_{x=d} \frac{w(x)}{y(x)} = 0$$

will be called a *focal solution* belonging to $x=d$. A focal solution belonging to $x=d$ will always exist, if $x=d$ is not its own first conjugate point on $(0, d]$. If $x=d$ is not its own first conjugate point on $(0, d]$, the first conjugate point of $x=d$ preceding $x=d$, if it exists, will be given by the first zero preceding $x=d$ of a focal solution belonging to $x=d$. These results and others follow from the results of §2 upon applying the transformation (10.3) to J.

If $[0, d]$ affords a minimum limit to J among curves (A, A^*)-admissible on $[0, d]$ and if b is any constant on $(0, d)$, $[0, b]$ will afford a minimum limit to J relative to curves A-admissible on $[0, b]$. We are thereby led to the following theorem.

THEOREM 10.1. *If $[0, d]$ affords a minimum limit to J among curves (A, A^*)-admissible on $[0, d]$, there can be no first conjugate point of $x=0$ on $[0, d)$ or first conjugate point of $x=d$ on $(0, d]$.*

The conclusions of the theorem are not independent, for it follows from Lemma 2.2 that if $x=0$ has no first conjugate point on $[0, d]$, $x=d$ will have no first conjugate point on $(0, d]$. Moreover the roles of $x=0$ and $x=d$ can be interchanged.

In §4 we have defined the singularity condition at $x=0$ belonging to $[0, b]$. The condition that

$$\liminf_{x=d} s[y(x), b] \leqq 0 \qquad (0 < b < d)$$

for each A^*-admissible curve for which

$$\liminf_{x=d} J(y)\Big|_b^x \neq +\infty$$

will be called the *singularity condition* at $x=d$ belonging to $[b, d]$. With this understood we have the following theorem.

THEOREM 10.2. *If $[0, d]$ affords a minimum limit to J, the singularity condition at $x=0$ belonging to $[0, b]$ and the singularity condition at $x=d$ belonging to $[b, d]$ must be satisfied by each curve $y=y(x)$ (A, A^*)-admissible on $[0, d]$ and each constant b on $(0, d)$.*

11. **Sufficient conditions.** If $u(x)$ is a solution of the E.E. which does not vanish on an interval $I=(0, d)$ and $y(x)$ is a function (A, A^*)-admissible on $[0, d]$, it will be convenient to set

$$E[x, y(x), y'(x), p[x, y(x)]] = \mathcal{E}[y(x), u(x)] \qquad \text{(on } I\text{)}, \tag{11.0}$$

where $p(x, y)$ is the slope of the field $y=cu(x)$ at the point (x, y). With this understood we state an extension of Theorem 3.1.

THEOREM 11.1. *If $y(x)$ is (A, A^*)-admissible on $[0, d]$ and $u(x)$ is a solution of the* E.E. *which does not vanish on $(0, d]$, then for $0<x_1<x_2<d$,*

$$J(y)\Big|_{x_1}^{x_2} = \int_{x_1}^{x_2} \mathcal{E}[y(x), u(x)]dx + S[y(x), u(x)]_{x_2}^{x_1}. \tag{11.1}$$

The proof of (11.1) is similar to the proof of (3.1). It is based on the formula

$$J(y)\Big|_{x_1}^{x_2} = \int_{x_1}^{x_2} \mathcal{E}dx + H_y\Big|_{x_1}^{x_2},$$

where H is the Hilbert integral set up for the field $y=cu(x)$. As in §3 one shows that

$$H_y\Big|_{x_1}^{x_2} = S[y(x), u(x)]_{x_2}^{x_1}$$

and (11.1) follows at once.

We are led to the following theorem.

THEOREM 11.2. *In order that $[0, d]$ afford a minimum limit to J, it is sufficient that there exist a solution $u(x)$ of the* E.E. *such that $u(x)\neq 0$ on $(0, d)$ and such that*

$$\liminf_{x_1=0,\, x_2=d} S[y(x), u(x)]_{x_2}^{x_1} \geqq 0$$

for every (A, A^*)-admissible curve $y=y(x)$.

Theorem 11.2 follows at once from (11.1).

An inequality. Let $z(x)$ and $\mu(x)$ be two solutions of the E.E. such that $z(b)=\mu(b)$ for some value of b on $(0, d)$ and such that $z(x)>0$ on $(0, b]$ and $\mu(x)>0$ on $[b, d)$. Let $y=y(x)$ be a curve which is (A, A^*)-admissible on $[0, d]$. By a proof similar to that of (11.1) we find that

$$J(y)\Big|_{x_1}^{b} = \int_{x_1}^{b} \mathcal{E}[y(x), z(x)]dx + S[y(x), z(x)]_{b}^{x_1}, \tag{11.2}$$

$$J(y)\Big|_{b}^{x_2} = \int_{b}^{x_2} \mathcal{E}[y(x), \mu(x)]dx + S[y(x), \mu(x)]_{x_2}^{b}. \tag{11.3}$$

We thus arrive at the following inequality

$$J(y)\Big|_{x_1}^{x_2} \geqq S[y(x), z(x)]_b^{x_1} + S[y(x), \mu(x)]_{x_2}^{b}. \tag{11.4}$$

We shall use (11.4) in proving the following theorem.

THEOREM 11.3. *If in the case $q(x) \equiv 0$ there exists a solution $u(x)$ of the* E.E. *for which*

$$u'(x) \leqq 0, \qquad u(x) > 0, \qquad (x > 0),$$

neighboring $x=0$ and a solution $v(x)$ such that

$$v'(x) \geqq 0, \qquad v(x) > 0, \qquad (x < d),$$

neighboring $x=d$, and if $x=0$ possesses no conjugate point on $[0, d)$, J assumes a minimum limit on $[0, d]$ among curves which are (A, A^*)-admissible.

Let $w(x)$ be a solution of the E.E. which is positive on $(0, d)$. Let b be a constant on $(0, d)$. Let $z(x)$ be a solution such that

$$z(b) = w(b), \qquad z'(b) = w'(b) - e \qquad (e > 0). \tag{11.5}$$

Note that

$$z(x) \geqq w(x) > 0 \qquad (0 < x \leqq b).$$

Similarly let $\mu(x)$ be a solution such that

$$\mu(b) = w(b), \qquad \mu'(b) = w'(b) + e. \tag{11.6}$$

Observe that

$$\mu(x) \geqq w(x) > 0 \qquad (b \leqq x \leqq d).$$

Let $y=y(x)$ be an (A, A^*)-admissible curve. We turn to the relation (11.4).

If the above constant e is sufficiently small, neither $z(x)$ nor $\mu(x)$ will be a focal solution belonging to $x=0$ or $x=d$. It follows as in the proof of Theorem 5.3 that

$$\liminf_{x_1=0} S[y(x), z(x)] \geqq 0, \tag{11.7}$$

$$\liminf_{x_2=d} S[y(x), \mu(x)] \leqq 0. \tag{11.8}$$

From (11.4) we see that

$$\liminf_{x_1=0,\, x_2=d} J(y)\Big|_{x_1}^{x_2} \geqq S[y(b), \mu(b)] - S[y(b), z(b)]. \tag{11.9}$$

Upon letting the constant e approach zero the right member of (11.9) tends to zero, and the theorem follows directly.

The following theorem is an easy consequence of the preceding (cf. proof of Theorem 5.4).

THEOREM 11.4. *If $p \geqq 0$ neighboring $x=0$ and $x=d$, if $q \equiv 0$, and if there is no point on $[0, d)$ conjugate to $x=0$, $[0, d]$ affords a minimum limit to J among (A, A^*)-admissible curves.*

12. **The regular case: two singular points.** We now suppose that the integrand f in (10.1) may be given the form of the integrand in (8.0) or (8.1) neighboring $x=0$, and that neighboring $x=d$ it may be given the form

$$(12.1)\qquad \begin{aligned}(x-d)^{\alpha^*}g^*(x)y'^2 - (x-d)^{\alpha^*-2}h^*(x)y^2 \\ \equiv (x-d)^{\alpha^*}g^*(x)y'^2 - (x-d)^{m^*-2}k^*(x)y^2,\end{aligned}$$

where $g^*(x)$, $h^*(x)$, and $k^*(x)$ are analytic in x neighboring $x=d$, with $g^*(x)>0$ and $k^*(d)\neq 0$. The E.E. corresponding to such a functional will have a regular singular point at $x=0$ and at $x=d$. The indicial roots corresponding to the singular point $x=d$ will be denoted by σ^*, and τ^*, with $\sigma^* \geqq \tau^*$ when σ^* and τ^* are real.

Our first theorem is as follows.

THEOREM 12.1. *Necessary and sufficient conditions that $[0, d]$ afford a minimum limit to J among curves which are (A, A^*)-admissible are that $[0, d]$ contain no point conjugate to $x=0$ and that $m>1$ if $k(0)>0$ and $m^*>1$ if $k^*(0)>0$.*

That the conditions of the theorem are necessary follows from the theorems for one singular point. One considers the segments $[0, b]$ and $[b, d]$ with b on $(0, d)$.

To prove the conditions sufficient we consider the limits

$$(12.2)\qquad \lim_{x_1=0} \int_{x_1}^{b} x^{\alpha}y'^2dx, \qquad \lim_{x_2=d} \int_{b}^{x_2} (x-d)^{\alpha^*}y'^2dx,$$

where $y(x)$ is (A, A^*)-admissible on $[0, d]$ and $(0<b<d)$. If one or both of the limits (12.2) are infinite, it is clear that

$$(12.3)\qquad \lim_{x_1=0,\, x_2=d} \inf J(y)\Big|_{x_1}^{x_2} > 0,$$

as desired. Suppose both of the limits (12.2) are finite. It follows from Lemma 8.1 that $S[y(x), u(x)]$ tends to zero as x tends to zero or d where $u(x)$ is any solution of the E.E. which is not zero on $(0, d)$.

The theorem follows from Theorem 11.2.

Theorem 12.1 has the following corollary.

COROLLARY. *If* $[0, d)$ *contains no point conjugate to* $x=0$, *necessary and sufficient conditions that* J *assume a minimum limit among* (A, A^*)*-admissible curves on* $[0, d]$ *are that for each constant* b *on* $(0, d)$ J *assumes a minimum limit on* $[0, b]$ *and* $[b, d]$ *respectively among curves* A*-admissible and* A^**-admissible on* $[0, b]$ *and* $[b, d]$ *respectively.*

We next consider the possibility of improper minimum limits. Theorem 8.2 includes the conditions

$$(12.4)' \qquad m > 1, \qquad \alpha < 1 \text{ if } k(0) > 0,$$

$$(12.4)'' \qquad \alpha = m, \text{ or } \alpha > 1 \text{ if } k(0) < 0.$$

We term these conditions the *finite conditions* in Theorem 8.2. Analogous finite conditions at the point $x=d$ are obtained by placing an asterisk above m, α, and k in (12.4). With this understood we state the following theorem.

THEOREM 12.2. *Necessary and sufficient conditions that* $[0, d]$ *afford an improper minimum limit to* J *among curves which are* (A, A^*)*-admissible on* $[0, d]$ *are* (1) *that a focal solution belonging to* $x=0$ *be a focal solution belonging to* $x=d$ *and* (2) *that the finite conditions* (12.4) *hold at* $x=0$ *and the analogous conditions hold at* $x=d$.

We begin by proving condition (1) necessary.

If $[0, d]$ affords an improper minimum limit to J, there exists an admissible curve $y=y(x)\not\equiv 0$ such that

$$(12.5) \qquad \liminf_{x_1=0,\, x_2=d} J(y) \Big|_{x_1}^{x_2} = 0.$$

Suppose in particular that $y(b)>0$ where $0<b<d$. Let $z(x)$ and $\mu(x)$ be focal solutions belonging to $x=0$ and $x=d$ respectively such that

$$(12.6) \qquad z(b) = \mu(b) = y(b).$$

If condition (1) is not necessary, as we now assume, $z(x)\not\equiv\mu(x)$. It follows that

$$(12.7) \qquad z'(b) > \mu'(b)$$

as one sees from the properties of focal solutions.

We shall make use of the relation (11.4). If (12.5) holds, it is impossible for either of the limits (12.2) to be infinite. It follows from Lemma 8.1, as we have noted previously, that (11.7) and (11.8) hold, and accordingly (11.9). Upon making use of the definitions of $z(x)$ and $\mu(x)$ and the relation (12.7), (11.9) takes the form

$$\liminf_{x_1=0,\, x_2=d} J(y) \Big|_{x_1}^{x_2} \geqq r(b)y(b)[z'(b) - \mu'(b)] > 0 \tag{12.8}$$

contrary to (12.5).

We infer that $z(x) \equiv \mu(x)$, and condition (1) of the theorem is proved necessary.

To prove conditions (2) necessary we trace (12.8) back to equations (11.2) and (11.3) and observe that the equality holds in (12.8) only if

$$y(x) \equiv z(x) \equiv \mu(x);$$

that is, if $y(x)$ is a focal solution belonging to $x=0$ and $x=d$. That conditions (2) are necessary and that conditions (1) and (2) are sufficient follow now as in the proof of Theorem 8.2.

It is clear that we can use theorems on minimum limits when the functional has one singular point to suggest theorems on minimum limits for the various types of (X, Y)-admissibility. The methods of proof will combine the methods of §9 and the present section. We give one additional theorem suggested by Theorems 9.1 and 9.2.

THEOREM 12.3. *Sufficient conditions that* $[0, d]$ *afford a minimum limit to* J *among curves which are* (α, F)*-admissible are that* $[0, d)$ *contain no point conjugate to* $x=0$ *and that* $\tau^*<0$ *or* $\sigma^*=\tau^*=0$. *This minimum limit is improper if and only if a focal solution belonging to* $x=0$ *is a focal solution belonging to* $x=d$, *if* $\sigma>0$, $\sigma-\tau>0$, *and* $\sigma^*>0$.

EXAMPLE 12.1. The following theorem is stated in H.L.P. It follows from Theorem 12.3.

If y' *belongs to* $L^2(0, \infty)$, $y_0=0$, *and* y *is not always zero, then*

$$J(y) = \int_0^\infty \left(4y'^2 - \frac{y^2}{x^2}\right) dx > 0. \tag{12.9}$$

The curves admitted by H.L.P. are a sub-class of those (α, F)-admissible on $[0, \infty]$. To apply the preceding theorem we make the transformation

$$t = \frac{x}{x+1}, \qquad x = \frac{t}{1-t}.$$

The transformed functional will have singular points at $t=0$ and $t=1$ with $\sigma=\tau=\frac{1}{2}$ and $\alpha^*=m^*=2$. Since

$$-1 = 1 - \alpha^* = \sigma^* + \tau^*,$$

we infer that $\tau^*<0$. The E.E. corresponding to (12.9) has the focal solution

$x^{1/2}$ so that the conjugate point condition is satisfied. The minimum limit afforded by $[0, \infty]$ is proper since $\sigma = \tau$.

BIBLIOGRAPHY

Bieberbach, Ludwig.
1. *Theorie der Differentialgleichungen*, Berlin, Springer, 1930.
Birkhoff, George D.
1. *A simplified treatment of the regular singular point*, these Transactions, vol. 11 (1910), pp. 199–202.
Bôcher, Maxime.
1. *Leçons sur les méthodes de Sturm*, Paris, Gauthier-Villars, 1917.
Hardy, G. H., Littlewood, J. E., Pólya, G.
1. *Inequalities*, Cambridge, University Press, 1934.
Kemble, E. C.
1. *Note on the Sturm-Liouville eigenvalue-eigenfunction problem with singular end-points*, Proceedings of the National Academy of Sciences, vol. 19 (1933), pp. 710–714.
de la Vallée Poussin, Ch.-J.
1. *Analyse Infinitésimale*, Tome I, Louvain, 1914.
Morse, Marston.
1. *The Calculus of Variations in the Large*, American Mathematical Society Colloquium Publications, vol. 18, New York, 1934.

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

SOME THEOREMS ON FOURIER TRANSFORMS AND CONJUGATE TRIGONOMETRIC INTEGRALS†

BY

R. P. BOAS, JR.

1. **Introduction.** Functions whose Fourier transforms vanish for large values of the argument have been studied by Paley and Wiener.‡ It is natural, then, to ask what properties are possessed by functions whose Fourier transforms vanish over a finite interval. The two classes of functions are evidently closely related, since a function whose Fourier transform vanishes on $(-A, A)$ is the difference of two suitably chosen functions, one of which has its Fourier transform vanishing outside $(-A, A)$. In this paper, however, we obtain, for a function having a Fourier transform vanishing over a finite interval, a criterion which does not seem to be trivially related to known results. It is stated in terms of a transform which is closely related to the Hilbert transform, and which may be written in the simplest (and probably most interesting) case,

$$f^*(x) = \frac{1}{\pi}\int_0^\infty \frac{f(x+t) - f(x-t)}{t^2} \sin t\, dt. \tag{1}$$

With this notation, we shall show that a necessary and sufficient condition for a function $f(x)$ belonging to $L^2(-\infty, \infty)$ to have a Fourier transform vanishing almost everywhere on $(-1, 1)$ is that $f^{**}(x) = -f(x)$ almost everywhere; $f^{**}(x)$ is the result of applying the transform (1) to $f^*(x)$. The result holds also if $(\sin t)/t$ is replaced in the integrand in (1) by a certain more general function $H(t)$.

As a preliminary result, we discuss the effect of the general transform $f^*(x)$ on a trigonometric Stieltjes integral,

$$f(x) = \int_{-R}^{R} e^{ixt} d\alpha(t), \tag{2}$$

where $\alpha(t)$ is a normalized§ complex-valued function of bounded variation. It is found that if $\alpha(t)$ is constant on $-1<t<0$ and on $0<t<1$, then $f^*(x)$ is equal to the conjugate trigonometric integral,

† Presented to the Society, April 10, 1936; received by the editors, January 10, 1936.

‡ R. E. A. C. Paley and N. Wiener, *Fourier Transforms in the Complex Domain*, American Mathematical Society Colloquium Publications, vol. 19, 1934, pp. 13, 24, 68 ff.

§ $\alpha(t)$ is said to be normalized if $\alpha(t) = \frac{1}{2}[\alpha(t+0)+\alpha(t-0)]$, $-R<t<R$.

$$\bar{f}(x) = i\int_{-R}^{R} e^{ixt}\,\operatorname{sgn} t\, d\alpha(t).\dagger$$

This includes the case where $f(x)$ is a trigonometric polynomial.‡

The relation $\bar{f}(x) = f^*(x)$, when valid, proves to be effective for establishing inequalities for $\bar{f}(x)$. For this application we shall need also generalizations of S. Bernstein's classical inequality for the derivative of a trigonometric polynomial;§ these we obtain in §4. For the sake of completeness it has seemed desirable to establish stronger results than we actually need, and which generalize to integrals of the form (2) extensions of Bernstein's theorem which have been given for trigonometric polynomials by G. Szegö‖ and by A. Zygmund.¶

Bernstein's theorem states that if

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{R} (a_n \cos nx + b_n \sin nx)$$

is a trigonometric polynomial of order R (with real coefficients), and if $|f(x)| \leqq M$, $-\infty < x < \infty$, then $|f'(x)| \leqq MR$, $-\infty < x < \infty$. Bernstein has shown†† that the same conclusion follows if $f(x)$ is merely an entire function with the property that $\limsup_{n\to\infty} |f^{(n)}(x)|^{1/n} = R$; this includes functions $f(x)$ of the form (2), and would be sufficient for our purposes.

When $f(x)$ is a trigonometric polynomial, Szegö's result is that $|f(x)| \leqq M$ implies not only Bernstein's inequality but also

$$|R\sigma(x) + \bar{f}'(x)\cos\theta + f'(x)\sin\theta| \leqq MR \tag{3}$$

for any real θ, where

$$\sigma(x) = \frac{a_0}{2} + \sum_{n=1}^{R-1}\left(1 - \frac{n}{R}\right)(a_n \cos nx + b_n \sin nx)$$

is an arithmetic mean of $f(x)$. An equivalent statement is

---

† $\operatorname{sgn} t = t/|t|$, $t \neq 0$; $\operatorname{sgn} 0 = 0$.

‡ Except that our definition of the conjugate yields the negative of the usual conjugate in the case of a trigonometric polynomial; our conjugate is so defined that it will coincide with the Hilbert transform.

§ I am indebted to Professor D. V. Widder for originally calling my attention to the question of generalizing Bernstein's theorem to trigonometric integrals.

‖ G. Szegö, *Über einen Satz des Herrn Serge Bernstein*, Schriften der Königsberger gelehrten Gesellschaft, vol. 5 (1928), pp. 59–70.

¶ A. Zygmund, *A remark on conjugate series*, Proceedings of the London Mathematical Society, (2), vol. 34 (1932), pp. 392–400.

†† S. Bernstein, *Leçons sur les Propriétés Extrémales et la Meilleure Approximation des Fonctions Analytiques d'une Variable Réelle*, 1926, p. 102.

$$R\,|\,\sigma(x)\,|\,+\,[(\bar{f}'(x))^2+(f'(x))^2]^{1/2}\leqq MR.$$

Zygmund showed that if $\phi(u)$ is a non-decreasing convex function of $u(u\geqq 0)$, then

$$\int_0^{2\pi}\phi\left(\frac{|\,f'(x)\,|}{R}\right)dx\leqq\int_0^{2\pi}\phi(\,|\,f(x)\,|\,)dx$$

with the same inequality for $\phi(|\,\bar{f}'(x)\,|/R)$.

These are the results of Szegö and Zygmund which we generalize. We show that if $f(x)$ has the form (2), with $\bar{\alpha}(t)=\alpha(-t)$,† and if $|f(x)|\leqq M$, $-\infty<x<\infty$, then (3) holds, where

$$\sigma(x)=\int_{-R}^{R}\left(1-\frac{|\,t\,|}{R}\right)e^{ixt}d\alpha(t);$$

and that Zygmund's result may be generalized to the inequality

$$\int_{-\infty}^{\infty}\phi\left(\left|\,\sigma(x)+\frac{\cos\theta}{R}\,\bar{f}'(x)+\frac{\sin\theta}{R}\,f'(x)\,\right|\right)dx\leqq\int_{-\infty}^{\infty}\phi(\,|f(x)\,|)dx,$$

where $\phi(u)$ is non-negative, non-decreasing, and convex.

From (3) we may also obtain the generalization to integrals of a result of van der Corput and Schaake,‡ namely, that under the same hypotheses on $f(x)$,

$$|\,f'(x)\,|\leqq R[M^2-(f(x))^2]^{1/2}.$$

In §5 we apply our results to obtain inequalities for conjugate trigonometric integrals. We can generalize to trigonometric integrals a theorem of M. Fekete,§ which states that if $f(x)$ is measurable on $(-\pi,\pi)$ and periodic with the period 2π; if $|f(x)|\leqq M$; and if $\bar{s}_n(x)$ is the sum of the first n terms of the series conjugate to the Fourier series of $f(x)$, then $|\,\bar{s}_n(x)|\leqq CM\log n$, where C is an absolute constant. We shall show that if $f(x)$ has the form (2), with $\bar{\alpha}(t)=\alpha(-t)$, and $\alpha(t)$ constant on $0<t<1$, then $|f(x)|\leqq M$, $-\infty<x<\infty$, implies

† Professor Szegö has pointed out to me that this last restriction may be removed by an elementary argument of a familiar type. The same remark applies to the inequality (4). (Added in proof.)

‡ J. G. van der Corput and G. Schaake, *Ungleichungen für Polynome und trigonometrische Polynome*, Compositio Mathematica, vol. 2 (1935), pp. 321–361; p. 337. The inequality is actually implicit in Szegö's inequality (3).

§ M. Fekete, *Über einen Satz des Herrn Serge Bernstein*, Journal für die reine und angewandte Mathematik, vol. 146 (1916), pp. 88–94; p. 91. Fekete gives no proof; the only proof which I have seen is that given by T. Carleman, *Über die Fourierkoeffizienten einer stetigen Funktion*, Acta Mathematica, vol. 41 (1918), pp. 377–384; p. 383. I am indebted to Professor G. H. Hardy for calling my attention to the problem of bounds for the conjugate of a trigonometric polynomial.

$$|\bar{f}(x)| < M\left(A + \frac{2}{\pi}\log R\right), \tag{4}$$

where A is an absolute constant $<4/\pi$. The constant $2/\pi$ multiplying $\log R$ is the best possible constant. Our result is less precise than Fekete's in that it requires a more restrictive hypothesis on $f(x)$; on the other hand, it applies to a wider class of functions, and fixes the best possible value of the constant C.

The analogy between the inequality (4) and S. Bernstein's inequality for derivatives suggests that Zygmund's generalization (mentioned above) of Bernstein's theorem to an inequality for a mean value of the derivative should have an analogue for the conjugate integral. We establish such an analogue in Theorem 8.

2. **A preliminary theorem.** We consider first trigonometric Stieltjes integrals over a finite range; that is, functions of the form

$$f(x) = \int_{-R}^{R} e^{ixt}d\alpha(t), \tag{2}$$

where $\alpha(t)$ is a normalized complex-valued function of the real variable t, of bounded variation on $(-R, R)$. There is no loss in generality from assuming $R>1$. Then we write

$$\begin{aligned} f_1(x) &= \int_{-1}^{1} e^{ixt}d\alpha(t), \\ f_2(x) &= \left(\int_{-R}^{-1} + \int_{1}^{R}\right) e^{ixt}d\alpha(t), \end{aligned} \tag{5}$$

so that $f(x)=f_1(x)+f_2(x)$.

We now introduce the notation, for any $f(x)$ defined on $(-\infty, \infty)$,

$$f^*(x) = \frac{1}{\pi}\int_0^{\infty} \frac{f(x+t)-f(x-t)}{t} H(t)dt, \tag{6}$$

for any x for which the integral exists in the sense of $\lim_{\epsilon\to 0, T\to\infty}\int_{\epsilon}^{T}$. $H(t)$ is a function satisfying Conditions A, below.

Let us denote by $\bar{f}(x)$ the Hilbert transform† of $f(x)$,

$$\bar{f}(x) = \frac{1}{\pi}\int_0^{\infty} \frac{f(x+t)-f(x-t)}{t}\,dt. \tag{7}$$

† We shall use the bar to denote both the Hilbert transform and the complex conjugate, but no confusion should result.

It is known that if $f(x)$ belongs to $L^2(-\infty, \infty)$, then $\bar{f}(x)$ exists for almost all x, and also belongs to $L^2(-\infty, \infty)$.†

We shall require $H(t)$ to satisfy the following conditions.

CONDITIONS A. 1. *$h(s)$ is a continuous function of bounded variation on $-1 \leqq s \leqq 1$, and*

$$H(t) = \int_{-1}^{1} \cos st \, dh(s).$$

2. *For whatever class of functions $f(x)$ is under consideration, $H(t)$ shall be such that the integral* (6) *converges with respect to the upper limit whenever* (7) *does so.*

The second condition will in particular be satisfied if $H(t)$ has bounded variation on (T, ∞) for some T; or if $h(s)$ is the integral of a function $k(s)$ of bounded variation on $(-1, 1)$, and the class of functions $f(x)$ is the class of functions bounded as $x \to \infty$.

We shall establish the following preliminary theorem.

THEOREM 1. *If $f(x)$ and $f_2(x)$ are defined by* (2) *and* (5), *and if $H(t)$ satisfies Conditions* A, *then $f^*(x)$ exists for all x, and*

$$f^*(x) = H(0)\bar{f}_2(x) + i\int_{-1}^{1} e^{ixt}[h(t) - h(-t)]\, d\alpha(t). \tag{8}$$

It will be sufficient to establish the theorem for the special case

$$f(x) = \int_{-R}^{R} \sin xt \, d\beta(t), \tag{9}$$

where $\beta(t)$ is a normalized function of bounded variation, and for the special point $x = 0$. In this case (8) becomes

$$f^*(0) = H(0)\bar{f}_2(0) + \int_{-1}^{1} [h(t) - h(-t)]\, d\beta(t).$$

To see this,‡ suppose Theorem 1 established in this special case, and set, with $f(x)$ now defined by (2),

† N. Lusin, *Sur la convergence des séries trigonométriques de Fourier*, Paris Comptes Rendus, vol. 156 (1913), pp. 1655–1658. For references to other proofs, see A. Zygmund, *Trigonometrical Series*, 1935, Chapter 7; and G. H. Hardy and J. E. Littlewood, *The allied series of a Fourier series*, Proceedings of the London Mathematical Society, (2), vol. 24 (1925), pp. 211–246.

‡ This device was suggested by a remark made by L. Fejér and by E. Landau, in connection with S. Bernstein's theorem for derivatives. Fekete, op. cit., pp. 88, 93, attributes the idea to Fejér; Bernstein, *Sur une propriété des fonctions entières*, Paris Comptes Rendus, vol. 176 (1923), pp. 1603–1605, attributes it to Landau.

$$g(y) = \frac{1}{2}\left[f(x+y) - f(x-y)\right] = \frac{1}{2}\int_{-R}^{R}\left[e^{i(x+y)t} - e^{i(x-y)t}\right]d\alpha(t),$$

$$(10) \qquad g(y) = \int_{-R}^{R} \sin yt\, d\beta(t), \qquad \beta(t) = \int_{-R}^{t} ie^{ixu}d\alpha(u).$$

Then $g(y)$ has the form (9), and

$$\begin{aligned} g^*(y) &= \frac{1}{\pi}\int_0^\infty \frac{g(y+t) - g(y-t)}{t} H(t)dt \\ &= \frac{1}{2\pi}\int_0^\infty \frac{f(x+y+t) - f(x+y-t)}{t} H(t)dt \\ &\quad + \frac{1}{2\pi}\int_0^\infty \frac{f(x-y+t) - f(x-y-t)}{t} H(t)dt \\ &= \tfrac{1}{2}\left[f^*(x+y) + f^*(x-y)\right]. \end{aligned}$$

Thus, $g^*(0) = f^*(x)$. Similarly, $\bar{g}(0) = \bar{f}(x)$. Finally,

$$\int_{-1}^{1}\left[h(t) - h(-t)\right]d\beta(t) = i\int_{-1}^{1} e^{ixt}\left[h(t) - h(-t)\right]d\alpha(t),$$

as we see by referring to (10).

We assume that the relation

$$(11) \qquad \bar{f}(x) = i\int_{-R}^{R} e^{ixt} \operatorname{sgn} t\, d\alpha(t)$$

for functions of the form (2), and in particular

$$(12) \qquad \bar{f}(x) = \int_{-R}^{R} \cos xt \operatorname{sgn} t\, d\beta(t)$$

for functions of the form (9), is known to hold for all x; it is easily established by substituting the explicit expression for $f(x)$ into (7), and changing the order of integration.†

We now show that if $f(x)$ is defined by (9), $f^*(0)$ exists and satisfies (8).

We have, since $f(x)$ is an odd function,

$$\bar{f}(0) = \frac{2}{\pi}\int_0^\infty \frac{f(t)}{t}\,dt,$$

$$f^*(0) = \frac{2}{\pi}\int_0^\infty \frac{f(t)H(t)}{t}\,dt,$$

† This fact is implicit in a number of places; but I do not know where to refer for an explicit statement.

provided that the integrals exist. The convergence of the first integral is already known; the second integral converges with respect to the upper limit, because of Conditions A, and its convergence at the lower limit follows from the existence of $f'(0)$.

Consider now

$$
\begin{aligned}
f^*(0) &= \frac{2}{\pi}\int_0^\infty \frac{f(t)}{t}\,H(t)dt \\
&= \frac{2}{\pi}\int_0^\infty \frac{dt}{t}\int_{-R}^{R}\sin tu\,d\alpha(u)\int_{-1}^{1}\cos st\,dh(s) \qquad (13)\\
&= \frac{2}{\pi}\int_{-R}^{R}d\alpha(u)\int_{-1}^{1}dh(s)\int_0^\infty \frac{\sin ut\cos st}{t}\,dt,
\end{aligned}
$$

provided that the change of the order of integration is justified. Assuming this for the present, we have

$$
\int_0^\infty \frac{\sin ut\cos st}{t}\,dt = \begin{cases} \frac{\pi}{2}\operatorname{sgn} u, & \left|\frac{s}{u}\right| < 1; \\ \frac{\pi}{4}\operatorname{sgn} u, & \left|\frac{s}{u}\right| = 1; \\ 0, & \left|\frac{s}{u}\right| > 1;\ s = u = 0; \end{cases}
$$

by a familiar formula. Thus,

$$
\begin{aligned}
f^*(0) &= \left(\int_{-R}^{-1} + \int_1^R\right)\operatorname{sgn} u\,d\alpha(u)\int_{-1}^{1}dh(s) + \int_{-1}^{1}\operatorname{sgn} u\,d\alpha(u)\int_{-|u|}^{|u|}dh(s) \\
&= H(0)\bar{f}_2(0) + \int_{-1}^{1}[h(u) - h(-u)]d\alpha(u)
\end{aligned}
$$

as desired. It remains to justify the change of the order of integration in (13). We have to show that

$$
\begin{aligned}
&\left(\int_0^T + \int_T^\infty\right)\frac{dt}{t}\int_{-R}^{R}\sin ut\,d\alpha(u)\int_{-1}^{1}\cos st\,dh(s) \qquad (14)\\
&\qquad = \int_{-R}^{R}d\alpha(u)\int_{-1}^{1}dh(s)\left(\int_0^T + \int_T^\infty\right)\frac{\sin ut\cos st}{t}\,dt.
\end{aligned}
$$

We note first that the improper integrals in (14) are actually convergent. Since the integrals over the finite part of the range are clearly equal, it is only necessary to show that $\lim_{T\to\infty} J(T) = 0$, where

$$J(T) = \int_T^\infty \frac{f(t)H(t)}{t}\,dt - \int_{-R}^{R} d\alpha(u) \int_{-1}^{1} dh(s) \int_T^\infty \frac{\sin ut \cos st}{t}\,dt$$
$$= J_1 - J_2.$$

We have

$$2J_2 = \int_{-R}^{R} d\alpha(u) \int_{-1}^{1} dh(s) \int_T^\infty \frac{\sin (u+s)t}{t}\,dt$$
$$+ \int_{-R}^{R} d\alpha(u) \int_{-1}^{1} dh(s) \int_T^\infty \frac{\sin (u-s)t}{t}\,dt$$
$$= J_3 + J_4.$$

It is a familiar fact in the theory of Fourier series that

$$\left| \int_a^b \frac{\sin xt}{t}\,dt \right| \leqq M, \tag{15}$$

where M is independent of x, a, and b; and that

$$\lim_{T\to+\infty} \int_{Tx}^\infty \frac{\sin v}{v}\,dv = 0, \tag{16}$$

uniformly for $x \geqq x_0 > 0$.

In J_3 the integral with respect to s over $(-u-\delta, -u+\delta)$ is $o(1)$ for $\delta \to 0$, since $h(s)$ is continuous; since the integral with respect to t is uniformly bounded, by (15); and since $\alpha(u)$ has bounded variation. (It is supposed that $h(s)$ is extended outside $(-1, 1)$, if necessary, so that it is continuous at ± 1, and constant outside $(-1, 1)$.) Similarly, in J_4 the integral with respect to s over $(u-\delta, u+\delta)$ is $o(1)$. Now, given any $\epsilon > 0$, we may take δ so small that these two terms are both less in absolute value than ϵ, and then fix δ. In the remaining part of J_3, we have $|u+s| \geqq \delta$, and

$$\int_T^\infty \frac{\sin (u+s)t}{t}\,dt = \pm \int_{T|u+s|}^\infty \frac{\sin v}{v}\,dv = o(1) \qquad (T \to \infty)$$

by (16). Similar reasoning applies to J_4.

Finally, J_1 is the remainder of a convergent integral, and hence $J_1 = o(1)$ as $T \to \infty$. Thus $\lim_{T\to\infty} J(T) = 0$, and Theorem 1 is fully established.

COROLLARY 1. *If $f(x)$ satisfies the hypotheses of Theorem* 1,

$$\frac{1}{\pi} \int_0^\infty \frac{f(x+t) - f(x-t)}{t^2} \sin t\,dt = -2f_1'(x) + \bar{f}_2(x).$$

This is the special case $h(s) = s$.

COROLLARY 2. *If $f(x)$ satisfies the hypotheses of Theorem* 1, *if* $H(0)=1$, *and if* $\alpha(t)\equiv$ *constant*, $-1<t<0$, *and* $0<t<1$, *then* $f^*(x)=\bar{f}(x)$.

This includes the case where $f(x)$ is a trigonometric polynomial,

$$f(x) = \frac{a_0}{2} + \sum_{k=0}^{n} (a_k \cos kx + b_k \sin kx).$$

COROLLARY 3. *If $f(x)$ satisfies the hypotheses of Theorem* 1, *and if* $H(0)=0$, *then*

$$f^*(x) = i\int_{-1}^{1} e^{ixt}[h(t) - h(-t)]d\alpha(t).$$

3. **Functions whose Fourier transforms vanish on $(-1, 1)$.** We shall need a result of E. C. Titchmarsh, which we state as a lemma.

LEMMA 1.† *If $f(x)$ belongs to $L^2(-\infty, \infty)$, and $F(x)$ is its Fourier transform, then the Fourier transform of $\bar{f}(x)$ is $E(x) = -iF(x)$ sgn x.*

We can now establish our main theorem.

THEOREM 2. *Let $f(x)$ belong to $L^2(-\infty, \infty)$; let $F(x)$ be the Fourier transform of $f(x)$. Let $H(t)$ satisfy Conditions* A, *and the further conditions that $H(0)=1$, and that $h(s)-h(-s)=\pm 1$ almost nowhere. A necessary and sufficient condition that $F(x)=0$, almost everywhere on $(-1, 1)$, is that*

$$f^{**}(x) = -f(x)$$

almost everywhere.

Here $f^*(x)$ is defined by (6), and $f^{**}(x)$ is related to $f^*(x)$ as $f^*(x)$ is to $f(x)$.

We establish first the necessity of the condition. We notice first that

$$G(t) = \frac{1}{t}(1 - H(t)) \tag{17}$$

is a function of L^2, and that its Fourier sine transform is

$$g(t) = \left(\frac{2}{\pi}\right)^{1/2}\int_0^\infty G(u) \sin ut\, du = \begin{cases} \left(\frac{\pi}{2}\right)^{1/2}(1 - h(t) + h(-t)), \\ \qquad\qquad -1 \leqq t \leqq 1, \\ 0, \quad |t| \geqq 1. \end{cases} \tag{18}$$

† E. C. Titchmarsh, *Conjugate trigonometric integrals*, Proceedings of the London Mathematical Society, (2), vol. 24 (1925), pp. 109–130; p. 114.

By Plancherel's theorem, for almost all x

$$f(x) = \frac{1}{(2\pi)^{1/2}} \text{l.i.m.}_{a\to\infty} \int_{-a}^{a} e^{-ixu}F(u)du. \tag{19}$$

Let now x be given and fixed. Then for almost all t, by (19),

$$\begin{aligned} k(t) \equiv f(x+t) - f(x-t) &= \frac{-2i}{(2\pi)^{1/2}} \text{l.i.m.}_{a\to\infty} \int_{-a}^{a} e^{-ixu}F(u) \sin tu\, du \\ &= \left(\frac{2}{\pi}\right)^{1/2} \text{l.i.m.}_{a\to\infty} \int_{0}^{a} \sin tu\, K(u)du, \end{aligned} \tag{20}$$

where $K(u) = -i[e^{-ixu}F(u) - e^{ixu}F(-u)]$, so that $K(u) = 0$ almost everywhere on $(0, 1)$.

By the Parseval relation for Fourier transforms,

$$\int_0^\infty k(y)G(y)dy = \int_0^\infty K(y)g(y)dy = 0, \tag{21}$$

since $g(y) \equiv 0$ on $(1, \infty)$, and $K(y) = 0$ almost everywhere on $(0, 1)$. Referring to (17) and (20), we see that (21) is equivalent to $\bar{f}(x) = f^*(x)$ whenever either of these functions exists; and $\bar{f}(x)$ exists almost everywhere.

Now $\bar{f}(x)$ belongs to $L^2(-\infty, \infty)$; let $E(x)$ be the Fourier transform of $\bar{f}(x)$. By Lemma 1, $E(x) = -iF(x)$ sgn x, so that $E(x) = 0$ almost everywhere on $(-1, 1)$. We may now apply our previous reasoning, with only formal modifications, to $\bar{f}(x)$, and find

$$-f(x) = \bar{f}(x) = f^{**}(x).$$

The necessity of the condition is thus established.

Now let $f(x)$ belonging to $L^2(-\infty, \infty)$ have $f^{**}(x) = -f(x)$. Set

$$k(x) = f(x) - \frac{1}{(2\pi)^{1/2}} \int_{-1}^{1} F(t)e^{-ixt}dt, \tag{22}$$

where $F(x)$ is the Fourier transform of $f(x)$. It is clear that $k(x)$ belongs to $L^2(-\infty, \infty)$; hence $k(x)$ has a Fourier transform $K(x)$. It is clear that $K(x) = 0$ almost everywhere on $(-1, 1)$. Hence, by what has already been proved, $k^{**}(x) = -k(x)$. But, by Theorem 1 and our hypothesis,

$$k^{**}(x) = -f(x) + \frac{1}{(2\pi)^{1/2}} \int_{-1}^{1} F(t)[h(t) - h(-t)]^2 e^{-ixt}dt.$$

Hence,

$$\int_{-1}^{1} F(t)(1 - [h(t) - h(-t)]^2)e^{-ixt}dt = 0$$

for almost all x. This is possible only if

$$F(t)(1 - [h(t) - h(-t)]^2) = 0$$

almost everywhere on $(-1, 1)$. And since almost everywhere $h(t) - h(-t) \neq \pm 1$, we have finally $F(t) = 0$, almost everywhere on $(-1, 1)$. This completes the proof of the theorem.

In Theorem 2, we assumed that $H(0) = 1$. This assumption was made to ensure that $H(0) \neq 0$, so that it was possible to divide by it. It is natural to ask what happens to Theorem 2 if we allow $H(0)$ to vanish. We establish the following theorem.

THEOREM 3. *Let $f(x)$ belong to $L^2(-\infty, \infty)$, and let $F(x)$ be the Fourier transform of $f(x)$. Let $H(t)$ satisfy Conditions* A; *let $H(0) = 0$, and let $h(s) = h(-s)$ almost nowhere. A necessary and sufficient condition that $F(x) = 0$ almost everywhere on $(-1, 1)$ is that $f^*(x) = 0$ almost everywhere.*

The proof is much the same as that of Theorem 3, but simplified because Lemma 1 is no longer needed.

To establish the necessity of the condition, we set now $G(t) = H(t)/t$. Since $H(0) = 0$, we find that the Fourier sine transform of $G(t)$ is

$$(23) \qquad g(t) = \left(\frac{2}{\pi}\right)^{1/2} \int_0^\infty G(u) \sin ut\, du = \begin{cases} \left(\frac{\pi}{2}\right)^{1/2} [h(t) - h(-t)], \\ \qquad\qquad -1 \leqq t \leqq 1, \\ 0, \quad |t| \geqq 1. \end{cases}$$

The relation (20) holds as before, and (21) holds also, if $g(t)$ is defined by (23). But now

$$\frac{1}{\pi} \int_0^\infty k(y)G(y)dy = f^*(x) = 0$$

almost everywhere.

On the other hand, if $f^*(x) = 0$ almost everywhere, we define $k(x)$ by (22), and $K(x)$ as before. Since $K(x)$ clearly vanishes almost everywhere on $(-1, 1)$, the first part of the proof gives

$$k^*(x) = \frac{-i}{(2\pi)^{1/2}} \int_{-1}^{1} F(t)\, [h(t) - h(-t)]e^{-ixt}dt = 0$$

for almost all x. This can happen, since $h(t) - h(-t)$ is almost nowhere zero, only if $F(t) = 0$ almost everywhere on $(-1, 1)$.

4. **Inequalities for the derivative of a trigonometric integral.** The generalizations of S. Bernstein's theorem, indicated in the introduction, may be conveniently formulated as follows.

THEOREM 4. *Let $f(z)$ be a Laplace integral of the form*

$$f(z) = \int_0^R e^{-zt}d\alpha(t) \quad (0 < R < \infty;\ z = x + iy), \tag{24}$$

where $\alpha(t)$ is a normalized complex-valued function of bounded variation. Let $|\Re[f(iy)]| \leqq M$, $-\infty < y < \infty$. Let θ be any real number. Then

$$|\Re[R\sigma(iy) + e^{i\theta}f'(iy)]| \leqq MR \qquad (-\infty < y < \infty), \tag{25}$$

where

$$\sigma(z) = \frac{1}{R}\int_0^R (R - t)e^{-zt}d\alpha(t)$$

is an "arithmetic mean" of $f(z)$.

Since (25) holds for all real θ, it yields as an equivalent statement

$$R|\sigma(iy)| + |f'(iy)| \leqq MR \qquad (-\infty < y < \infty),$$

and as somewhat weaker consequences

$$\Im[f'(iy)] \leqq R[M^2 - (\Re[f(iy)])^2]^{1/2} \qquad (-\infty < y < \infty), \tag{26}$$

$$(\Re[Rf(iy) - f'(iy)])^2 + 4(\Im[f'(iy)])^2 \leqq 4R^2M^2. \tag{27}$$

Relation (26) follows simply from (25) by use of the formula $R\sigma(z) = Rf(z) + f'(z)$.

THEOREM 5. *Let $f(z)$ have the form (24). Let $\phi(u)$ be a non-negative, non-decreasing convex function of $u (u \geqq 0)$. Let E denote $(-\infty, \infty)$ in general, and $(0, 2\pi)$ if $f(iy)$ has the period 2π (that is, if $f(z)$ reduces to a polynomial $g(w)$ in $w = e^{-z}$). Then for any real θ*

$$\int_E \phi\left(\left|\Re\left[\sigma(iy) + \frac{e^{i\theta}}{R}f'(iy)\right]\right|\right)dy \leqq \int_E \phi(|\Re[f(iy)]|)dy, \tag{28}$$

where the integral on the right may be infinite.

Similar integral inequalities corresponding to (26) and (27) may be easily formulated. When $E = (0, 2\pi)$, it is not necessary to require $\phi(u) \geqq 0$.

The results of Theorems 4 and 5 take on a more familiar form if we notice that if $f(z)$ has the form (24), $F(y) = \Re[f(iy)]$ has the form (2) (at least if $\Im(\alpha(t))$ is continuous at $t = 0$), and that then $\overline{F}(y) = -\Im[f(iy)]$, $F'(y) = -\Im[f'(iy)]$, $\overline{F}'(y) = -\Re[f'(iy)]$;

$$
\begin{aligned}
&-\Re[e^{i\theta}f'(iy)] = \overline{F}'(y)\cos\theta - F'(y)\sin\theta,\\
&(\Re[Rf(iy) - f'(iy)])^2 + 4(\Im[f'(iy)])^2 = (RF(y) + \overline{F}'(y))^2 + 4(F'(y))^2.
\end{aligned}
\tag{29}
$$

This allows us to put the results of Theorems 4 and 5 into the form in which they were stated in the introduction. Various other formulations are also possible.

If $f(z)$ is a polynomial in e^{-z}, of degree R, (25) contains for $\theta=0$ the original theorem of S. Bernstein, and for general θ reduces to Szegö's† theorem; (26) becomes a theorem of van der Corput and Schaake,† a consequence of Szegö's theorem; and if $\theta=0$ or $\pi/2$, (28) contains two theorems of Zygmund.† If in (24) $\alpha(t)$ is a step-function (with a finite or infinite number of jumps), (25) reduces essentially to a theorem of M. Fekete.‡

We establish Theorems 4 and 5 together, using a modification of the method of F. Riesz.§

We begin with $f(z)$ of the form (24) with $\alpha(t)$ continuous at $t=R$. It is convenient to require that $\alpha(0)=0$; this additional normalization has no effect on $f(z)$. We have, by a familiar inversion formula,‖

$$
\beta(t) \equiv \frac{1}{2\pi i}\int_{-i\infty}^{i\infty}\frac{f(s)}{s}e^{st}ds = \begin{cases} \alpha(t), & 0 < t \leqq R;\\ \alpha(R), & t \geqq R;\\ 0, & t < 0;\\ \tfrac{1}{2}\alpha(0+), & t = 0; \end{cases}
$$

the integral being taken in the Cauchy sense at $s=0$ and $s=\infty$. Then,

$$
\begin{aligned}
-f'(0) &= \int_0^R t\,d\alpha(t) = R\alpha(R) - \int_0^R \alpha(t)dt\\
&= R\alpha(R) - \frac{1}{2\pi i}\int_0^R dt\int_{-i\infty}^{i\infty}\frac{f(s)}{s}e^{st}ds\\
&= R\alpha(R) - \frac{1}{2\pi i}\int_0^R dt\int_{-\infty}^{\infty}\frac{f(iy)}{y}e^{iyt}dy\\
-f'(0) &= R\alpha(R) + \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{f(iy)}{y^2}(e^{iyR}-1)dy.
\end{aligned}
\tag{30}
$$

† For references, see footnotes on pp. 288, 289.

‡ According to Szegö, loc. cit. See also G. Szegö, *Zur Theorie der fastperiodischen Funktionen*, Mathematische Annalen, vol. 96 (1927), pp. 378–382; and the theorem of S. Bernstein referred to in Theorem 6.

§ F. Riesz, *Sur les polynômes trigonométriques*, Paris Comptes Rendus, vol. 158 (1914), pp. 1657–1661. Similar modifications of the same method have been used by Szegö and by Zygmund.

‖ See, e.g., D. V. Widder, *A generalization of Dirichlet's series and of Laplace's integrals by means of a Stieltjes integral*, these Transactions, vol. 31 (1929), pp. 694–743; p. 708. The result given there needs slight modifications to cover the case in hand.

The change of order of integration is justified without difficulty. We have also, in the same way,

$$
\begin{aligned}
(31)\qquad 0 &= \int_0^R t\, d\beta(2R - t) \\
&= R\alpha(R) + \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{f(iy)}{y^2} e^{2iRy}(1 - e^{-iRy})dy; \\
0 &= \int_0^R t\, d\beta(-t) \\
&= -\frac{1}{2\pi i}\int_0^R dt \int_{-i\infty}^{i\infty} \frac{f(s)}{s} e^{-st} ds
\end{aligned}
$$

and hence,

$$
(32)\qquad 0 = -\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\overline{f(iy)}}{y^2}(e^{iRy} - 1)dy,
$$

where $\overline{f(s)}$ is the complex conjugate of $f(s)$. Similarly,

$$
(33)\qquad 0 = \int_0^R t\, d\beta(t - 2R) = -\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\overline{f(iy)}}{y^2} e^{2iRy}(1 - e^{-iRy})dy.
$$

We subtract (30) from (31) and (32) from (33). This gives

$$
\begin{aligned}
f'(0) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{f(iy)}{y^2}(e^{iRy} - 1)^2 dy, \\
0 &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\overline{f(iy)}}{y^2}(e^{iRy} - 1)^2 dy;
\end{aligned}
$$

and hence, on adding the last two relations,

$$
(34)\qquad e^{i\theta}f'(0) = \frac{2}{\pi}\int_{-\infty}^{\infty} \Re[f(iy)]e^{iRy+i\theta}\frac{1 - \cos Ry}{y^2} dy.
$$

We take the conjugate of both sides of (34), and add the resulting relation to (34), obtaining

$$
(35)\qquad \Re[e^{i\theta}f'(0)] = \frac{2}{\pi}\int_{-\infty}^{\infty} \Re[f(iy)]\cos(\theta + Ry)\frac{1 - \cos Ry}{y^2} dy.
$$

In (35) we may replace $\cos(\theta+Ry)$ by $g(\theta+Ry)$, where $g(u)$ is any integrable function with the period 2π, having a Fourier series of the form

$$g(u) \sim \cos u + \sum_{n=2}^{\infty} a_n \cos nu.$$

For, reasoning similar to that used in deriving (35) shows that

$$\int_{-\infty}^{\infty} \Re[f(iy)] e^{ikRy} \frac{1 - \cos Ry}{y^2} dy = 0 \qquad (k = \pm 2, \pm 3, \cdots), \tag{36}$$

and it is permissible to substitute the Fourier series of $g(u)$ in the integral and integrate term by term.†

Now, for any y_0, $E(z) = f(iy_0 + z)$ has the form (24), and $E'(0) = f'(iy_0)$, $E(iy) = f(iy_0 + iy)$. Hence we have, introducing $g(\theta + Ry)$ in (35),

$$\Re[e^{i\theta} f'(iy_0)] = \frac{2}{\pi} \int_{-\infty}^{\infty} \Re[f(iy + iy_0)] g(\theta + Ry) \frac{1 - \cos Ry}{y^2} dy. \tag{37}$$

Now, we have

$$\sigma(iy) = \frac{1}{R} \int_0^R (R - t) e^{-iyt} d\alpha(t), \tag{38}$$

and by reasoning similar to that used in obtaining (37), we find

$$\Re[\sigma(iy_0)] = \frac{1}{\pi R} \int_{-\infty}^{\infty} \Re[f(iy + iy_0)] \frac{1 - \cos Ry}{y^2} dy.$$

Combining this with (37), we obtain

$$\begin{aligned} G(y_0) &\equiv |\, \Re[R\sigma(iy_0) + e^{i\theta} f'(iy_0)]\,| \\ &\leqq \frac{1}{\pi} \int_{-\infty}^{\infty} |\, \Re[f(iy + iy_0)]\,| \cdot |\, 1 + 2g(\theta + Ry)\,| \frac{1 - \cos Ry}{y^2} dy. \end{aligned} \tag{39}$$

We now specialize $g(u)$, taking a function nearly the same as that used by Szegö. Set

$$1 + 2g(u) = \frac{1}{2r} \left\{ \frac{(1 + r)^2(1 - r)}{1 - 2r \cos u + r^2} - \frac{(1 - r)^2(1 + r)}{1 + 2r \cos u + r^2} \right\} \equiv h(u) \tag{40}$$

$$(0 < r < 1).$$

This is readily seen to have the desired form, since

$$\frac{1 - r^2}{1 - 2r \cos u + r^2} = 1 + 2 \sum_{n=1}^{\infty} r^n \cos nu \qquad (-1 < r < 1).$$

† See, e.g., E. W. Hobson, *The Theory of Functions of a Real Variable and the Theory of Fourier's Series*, vol. 2, 1926, p. 583.

Then,

$$|1+2g(u)| \leqq k(u) = \frac{1}{2r}\left\{\frac{(1+r)^2(1-r)}{1-2r\cos u+r^2}+\frac{(1-r)^2(1+r)}{1+2r\cos u+r^2}\right\}.$$

The function $k(u)$ has a Fourier series of the form

$$k(u) = \frac{1}{r} + \sum_{n=1}^{\infty} c_n \cos nu.$$

It is easily verified that

$$\int_{-\infty}^{\infty} \cos n(\theta+Ry)\frac{1-\cos Ry}{y^2}\,dy = 0 \quad (n = 1, 2, \cdots),$$

so that

$$\frac{1}{\pi}\int_{-\infty}^{\infty} |h(\theta+Ry)|\frac{1-\cos Ry}{y^2}\,dy \leqq \frac{1}{\pi}\int_{-\infty}^{\infty} k(\theta+Ry)\frac{1-\cos Ry}{y^2}\,dy = \frac{R}{r}. \tag{41}$$

We apply this result to (39). We have in the first place, since by hypothesis $|\Re[f(iy+iy_0)]| \leqq M$,

$$G(y_0) \leqq \frac{MR}{r},$$

and, letting $r\to 1$,

$$G(y_0) \leqq RM.$$

Referring to (41), we obtain (25) of Theorem 4, established now whenever $\alpha(t)$ is continuous at $t=R$. If $\alpha(t)$ is discontinuous at $t=R$, take $R'>R$, and write

$$f(z) = \int_0^{R'} e^{-zt}\,d\alpha_1(t), \qquad \alpha_1(t) = \begin{cases}\alpha(t), & t \leqq R;\\ \alpha(R), & t \geqq R.\end{cases}$$

Then $\alpha_1(t)$ is continuous at $t=R'$, and we deduce

$$G(y_0) \leqq R'M.$$

Letting $R'\to R$, we have Theorem 4 fully established.

To obtain Theorem 5, we introduce the function $k(\theta+Ry)$ in (39), apply "Jensen's inequality" for integrals† to the right-hand side, and integrate over E. Then,

† G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, 1934, pp. 150–152.

$$\int_E \phi\left(\frac{r}{R}G(y_0)\right)dy_0$$

$$\leqq \frac{r}{R\pi}\int_{-\infty}^{\infty} k(\theta + Ry)\,\frac{1-\cos Ry}{y^2}\,dy \int_E \phi(|\,\Re[f(iy+iy_0)]\,|\,)dy_0.$$

In the integral over E on the right, apply the periodicity of $f(iy)$ if $E=(0, 2\pi)$, and make an obvious change of variable if $E=(-\infty, \infty)$; then

$$\int_E \phi\left(\frac{r}{R}G(y_0)\right)dy_0 \leqq \int_E \phi(|\,\Re[f(iy_0)]\,|\,)dy_0. \tag{42}$$

Finally, let $r\to 1$ in (42). We may take the limit under the integral sign, and we obtain the conclusions of Theorem 5, in the case where $\alpha(t)$ is continuous at $t=R$. This condition may be removed as it was in Theorem 4.

There is also a generalization of Bernstein's theorem along quite different lines, given by Bernstein himself. Zygmund's generalization of the original theorem suggests the following further generalization.

THEOREM 6. *Let $F(z)$ be an entire function such that*

$$\limsup_{n\to\infty} |F^{(n)}(z)|^{1/n} = r > 0$$

for some given z (hence for every z). Let $\phi(u)$ have the same properties as in Theorem 5; let E denote $(0, p)$ if $F(x)$ is a periodic function with period p, and $(-\infty, \infty)$ otherwise. Then

$$\int_E \phi\left(\frac{|F'(x)|}{r}\right)dx \leqq \int_E \phi(|F(x)|)dx, \tag{43}$$

where the integral on the right may be infinite.

The theorem of S. Bernstein† states that if $F(z)$ satisfies the hypotheses of Theorem 6, and if $|F(x)|\leqq M$, $-\infty<x<\infty$, then $|F'(x)|\leqq Mr$, $-\infty<x<\infty$. The proof of this theorem given by Pólya and Szegö‡ may easily be modified to give a proof of Theorem 6. A function $f(x)$ of the form (2) is a function of the class considered here, with $r=R$.

5. **Inequalities for the conjugate of a trigonometric integral.** We now apply Theorem 1, and the results of §4, to obtain the generalized theorem of Fekete, mentioned in the introduction.

---

† S. Bernstein, loc. cit.

‡ G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, 1925, vol. 2, section IV, problem 201, pp. 35, 219.

THEOREM 7. *If $f(x)$ is a trigonometric Stieltjes integral of the special form*

$$f(x) = c + \left(\int_{-R}^{-r} + \int_{r}^{R} \right) e^{ixt} d\alpha(t) \qquad (0 < r < R), \tag{44}$$

where $\alpha(t)$ is a normalized, complex-valued function of bounded variation with $\bar{\alpha}(t) = \alpha(-t)$; and if $|f(x)| \leqq M$, $-\infty < x < \infty$, then

$$|\bar{f}(x)| \leqq \left(A + \frac{2}{\pi} \log \frac{R}{r} \right) M, \tag{45}$$

where A is an absolute constant ($< 4/\pi$), independent of R, r, and $f(x)$.

$\bar{f}(x)$ is defined by (11) or (7), which are here equivalent. The factor $2/\pi$ is the best possible in a certain sense to be specified later.

If $\alpha(t)$ is a step-function with jumps only at integral points, $f(x)$ is a trigonometric polynomial, and this theorem, as we stated in the introduction, becomes in one respect more precise, in another respect less precise, than a theorem of Fekete.†

We establish Theorem 7 for $r = 1$; the general case follows by a change of variable.

By Corollaries 1 and 2 of Theorem 1,

$$\bar{f}(x) = \frac{1}{\pi} \int_0^\infty \frac{f(x+t) - f(x-t)}{t^2} \sin t \, dt,$$

$$\pi |\bar{f}(x)| \leqq \left| \int_0^{1/R} \frac{f(x+t) - f(x-t)}{t^2} \sin t \, dt \right| + \left| \int_{1/R}^{1} \right| + \left| \int_1^\infty \right|.$$

Applying Theorem 4, and the elementary fact that $[f(x+t) - f(x)]/t$ and $f'(x)$ have the same upper and lower bounds,

$$\left| \int_0^{1/R} \frac{f(x+t) - f(x-t)}{t^2} \sin t \, dt \right| \leqq 2M,$$

since $|(\sin t)/t| \leqq 1$. Then,

$$\left| \int_{1/R}^{1} \frac{f(x+t) - f(x-t)}{t^2} \sin t \, dt \right| \leqq 2M \int_{1/R}^{1} \frac{dt}{t} = 2M \log R,$$

$$\left| \int_1^\infty \frac{f(x+t) - f(x-t)}{t^2} \sin t \, dt \right| \leqq 2M \int_1^\infty \frac{|\sin t|}{t^2} \, dt$$

† M. Fekete, loc. cit.

$$< 2M \int_1^\infty \frac{dt}{t^2} = 2M,$$

$$|\bar{f}(x)| < \frac{2}{\pi} M(2 + \log R).$$

The constant $2/\pi$ multiplying $\log (R/r)$ is the best possible constant in the sense that for no constant $B<2/\pi$ is it true that

$$\limsup_{R\to\infty} \frac{\max |\bar{f}(x)|}{\log \frac{R}{r}} \leqq BM$$

for all $f(x)$ of the form (44), satisfying $|f(x)| \leqq M$. This may be seen by taking for $f(x)$ first order arithmetic means of the partial sums of the series $\sum_1^\infty (1/n) \sin nx$.

The value $4/\pi$ is certainly not the best possible value of the constant A; it is easy to improve it slightly by a more accurate estimate of $\int_1^\infty |\sin t| t^{-2} dt$.

It is clear that if $r=0$, we can say nothing about the order of magnitude of $\bar{f}(x)$. A simple example is furnished by

$$f(x) = \int_0^1 t^{\delta-1} \sin xt \, dt \qquad (0 < \delta < \tfrac{1}{2}).$$

It is perhaps worth while to remark that since $\bar{f}(x)$ is independent of the constant c in (44), we may make Theorem 7 somewhat sharper by replacing in (45) M by $\frac{1}{2}(k+K)$, where $-k \leqq f(x) \leqq K$.

We turn now to an inequality for a mean value of the conjugate of a trigonometric integral. Applying Theorem 5 (or Theorem 6), we shall prove the following theorem.

THEOREM 8. *Let $f(x)$ be* (A) *a trigonometric polynomial of order R, or* (B) *a trigonometric integral of the form*

$$f(x) = c + \left(\int_{-R}^{-1} + \int_1^R \right) e^{ixt} d\alpha(t) \qquad (R > 1), \tag{46}$$

where $\alpha(t)$ has the same properties as in Theorem 7. Let $\phi(u)$ be a non-negative, non-decreasing convex function of $u (u \geqq 0)$. Let E denote $(0, 2\pi)$ in case (A), *and $(-\infty, \infty)$ in case* (B). *Then*

$$\int_E \phi\left(\frac{\pi |\bar{f}(x)|}{4A(R)} \right) dx \leqq \frac{1}{2} \int_E [\phi(2|f(x)|) + \phi(|f(x)|)] dx, \tag{47}$$

where

$$A(R) = \int_{1/R}^{\infty} \frac{|\sin t|}{t^2}\, dt,$$

and the integral in the right-hand member of (47) *may be infinite.*

If $R \geqq 2$, we may replace the right-hand side of (47) by

$$\int_E \phi(|f(x)|)dx.$$

We have $(2/\pi)\log R + \frac{1}{2} < A(R) < \log R + 1$; this evaluation of $A(R)$ brings out the analogy between (47) and (43).

For the case $\phi(y) = y^r$, $r > 1$, a stronger result than case (A) of Theorem 8 is known. M. Riesz† has shown that for any $f(x)$ of $L^r(0, 2\pi)$, $r > 1$ (and hence for $f(x)$ a trigonometric polynomial), there is a constant A_r depending only on r such that

$$\int_0^{2\pi} |\bar{f}(x)|^r dx \leqq A_r{}^r \int_0^{2\pi} |f(x)|^r dx.$$

On the other hand, our result is still available when $\phi(y) = y$, while the result of Riesz breaks down for $r = 1$.

We require an additional lemma, which is a generalization of the fact that the derivative and the difference quotient of a continuous function have the same upper and lower bounds in an interval.

LEMMA 3. *Let $f(x)$ be absolutely continuous in $a \leqq x \leqq b$. Let $\phi(y)$ be non-negative, non-decreasing, and convex for $y \geqq 0$. Then*

$$\int_a^b \phi(|f'(x)|)dx = M \tag{48}$$

implies

$$\int_c^d \phi\left(\frac{|f(x+t) - f(x-t)|}{2t}\right)dx \leqq M, \tag{49}$$

provided that $0 \leqq t < \frac{1}{2}(b-a)$, $c - a > t$, $b - d > t$.

For the proof we have

$$J \equiv \int_c^d \phi\left(\frac{1}{2t}|f(x+t) - f(x-t)|\right)dx = \int_c^d \phi\left(\left|\int_{x-t}^{x+t} \frac{f'(u)}{2t}\, du\right|\right)dx$$

† See A. Zygmund, *Trigonometrical Series*, 1935, p. 147. E. C. Titchmarsh has given a similar result when $\phi(y)$ involves also logarithmic factors: *A theorem on conjugate functions*, Journal of the London Mathematical Society, vol. 5 (1930), pp. 88–91.

$$\leqq \frac{1}{2t}\int_c^d dx \int_{x-t}^{x+t} \phi(|f'(u)|)du,$$

by "Jensen's inequality." Then, integrating in the reverse order,

$$J \leqq \frac{1}{2t}\int_{c+t}^{d-t} \phi(|f'(u)|)du \int_{u-t}^{u+t} dx + \frac{1}{2t}\int_{c-t}^{c+t} \phi(|f'(u)|)du \int_c^{u+t} dx + \frac{1}{2t}\int_{d-t}^{d+t} \phi(|f'(u)|)du \int_{u-t}^{d} dx$$

$$\leqq \int_{c-t}^{d+t} \phi(|f'(u)|)du \leqq M.$$

COROLLARY 1. *Lemma 3 remains valid if a or b is infinite, or if both are, with any finite* $t \geqq 0$.

COROLLARY 2. *If* $f(x)$ *has the period* 2π, *the conclusion of Lemma 3 holds with* $a=c=0$, $b=d=2\pi$, $0 \leqq t \leqq 2\pi$.

This follows from a more precise evaluation of the integrals in the last step of the proof.

We are now in a position to establish Theorem 8. We have

$$\pi|\bar{f}(x)| \leqq \int_0^{1/R} \frac{|f(x+t)-f(x-t)|}{t}dt + \int_{1/R}^{\infty} \frac{|f(x+t)-f(x-t)|}{t^2}|\sin t|\,dt,$$

as in the proof of Theorem 7. Applying the convexity of $\phi(y)$,

$$2\phi\left(\frac{\pi|\bar{f}(x)|}{4A(R)}\right) \leqq \phi\left(\frac{1}{A(R)}\int_0^{1/R} \frac{|f(x+t)-f(x-t)|}{2t}dt\right) + \phi\left(\frac{1}{A(R)}\int_{1/R}^{\infty} \frac{|f(x+t)-f(x-t)|}{2t^2}|\sin t|\,dt\right).$$

To the two terms on the right we apply Jensen's inequality. This gives

$$2\phi\left(\frac{\pi|\bar{f}(x)|}{4A(R)}\right) \leqq R\int_0^{1/R} \phi\left(\frac{|f(x+t)-f(x-t)|}{2RtA(R)}\right)dt + \frac{1}{A(R)}\int_{1/R}^{\infty} \phi\left(\frac{|f(x+t)-f(x-t)|}{2}\right)\frac{|\sin t|}{t^2}dt.$$

Now, $\phi(y)$ is an increasing function. In the first term on the right, we use the fact that $1/A(R)<2$; in the second term, we use the inequality

$|f(x+t)-f(x-t)| \leqq |f(x+t)| + |f(x-t)|$, and apply the convexity of $\phi(y)$. This gives us

$$2\phi\left(\frac{\pi|\bar{f}(x)|}{4A(R)}\right) \leqq R\int_0^{1/R} \phi\left(\frac{|f(x+t)-f(x-t)|}{Rt}\right)dt + \frac{1}{2A(R)}\int_{1/R}^{\infty} [\phi(|f(x+t)|) + \phi(|f(x-t)|)]\frac{|\sin t|}{t^2}\,dt. \tag{50}$$

Now we integrate both sides of (50) over E, and change the order of integration, obtaining

$$2\int_E \phi\left(\frac{\pi|\bar{f}(x)|}{4A(R)}\right)dx \leqq R\int_0^{1/R} dt\int_E \phi\left(\frac{|f(x+t)-f(x-t)|}{Rt}\right)dx + \frac{1}{2A(R)}\int_{1/R}^{\infty}\frac{|\sin t|}{t^2}\,dt\int_E [\phi(|f(x+t)|) + \phi(|f(x-t)|)]dx. \tag{51}$$

Apply the appropriate corollary of Lemma 3 to the first term on the right of (51); in the second term, apply the periodicity of $f(x)$ in case (A), and make an obvious change of variable in case (B); then,

$$\begin{aligned} 2\int_E \phi\left(\frac{\pi|\bar{f}(x)|}{4A(R)}\right)dx &\leqq R\int_0^{1/R} dt\int_E \phi\left(\frac{2|f'(x)|}{R}\right)dx \\ &\quad + \frac{1}{A(R)}\int_{1/R}^{\infty}\frac{|\sin t|}{t^2}\,dt\int_E \phi(|f(x)|)dx \\ &\leqq \int_E \phi(2|f(x)|)dx + \int_E \phi(|f(x)|)dx, \end{aligned}$$

by Theorem 5 or 6. This completes the proof. It has been tacitly assumed that the integral in the right-hand member of (47) is finite; if it is infinite, the theorem is trivial.

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

DIFFERENTIABLE FUNCTIONS DEFINED IN ARBITRARY SUBSETS OF EUCLIDEAN SPACE*

BY

HASSLER WHITNEY

1. **Introduction.** In a former paper† we studied the differentiability of a function defined in closed subsets of Euclidean n-space E. We consider here the differentiability "about" an arbitrary point of a function defined in an arbitrary subset of E. We show in Theorem 1 that any function defined in a subset A of E which is differentiable about a subset B of E may be extended over E so that it remains differentiable about B. This theorem is a generalization of AE Lemma 2. We show further that any function of class C^m about a set B is of class C^{m-1} about an open set B' containing B. In the second part of the paper we consider some elementary properties of differentiable functions, such as: the sum or product of two such functions is such a function.‡ We end with the theorem that differentiability is a local property.§

2. **Definitions and elementary properties.** We use a one-dimensional notation as in AE. Thus $f_k(x)=f_{k_1\cdots k_n}(x_1,\cdots,x_n)$, $x^l=x_1^{l_1}\cdots x_n^{l_n}$, $l!=l_1!\cdots l_n!$, $D_kf(x)=\partial^{k_1+\cdots+k_n}f(x)/\partial x_1^{k_1}\cdots\partial x_n^{k_n}$, etc.; we set $\sigma_k=k_1+\cdots+k_n$, $r_{xy}=$distance from x to y. We always set $f(x)=f_0(x)$. Suppose the functions $f_k(x)$ for $\sigma_k\leqq m$ are defined in the subset A of Euclidean n-space E. Define $R_k(x';x)$ for x, x' in A by

---

* Presented to the Society, January 2, 1936; received by the editors October 26, 1935.

† *Analytic extensions of differentiable functions defined in closed sets*, these Transactions, vol. 36 (1934), pp. 63–89. We refer to this paper as AE. See also *Functions differentiable on the boundaries of regions*, Annals of Mathematics, vol. 35 (1934), pp. 482–485, and *Differentiable functions defined in closed sets*, I, these Transactions, vol. 36 (1934), pp. 369–387, which we refer to as F and D respectively.

P. Franklin in Theorem 1 of a paper *Derivatives of higher order as single limits*, Bulletin of the American Mathematical Society, vol. 41 (1935), pp. 573–582, has given a necessary and sufficient condition for the existence of a continuous mth derivative. We remark that this theorem is exactly the special case of Theorem I of D obtained by letting $f(x)$ be defined in an interval. It is also a special case of Theorem 2 of the author's *Derivatives, difference quotients, and Taylor's formula*, Bulletin of the American Mathematical Society, vol. 40 (1934), pp. 89–94 (see also Errata, p. 894). For his assumption is easily seen to imply the needed uniformity condition; it also implies at once that $f(x)$ is continuous, so that no considerations of measurability are necessary. His Theorem 2 should be compared with Theorems II and III of D.

‡ If the set is closed, these theorems may be proved by first extending the functions throughout E.

§ For the case of one variable this follows from D, Theorem I.

$$(1) \qquad f_k(x') = \sum_{\sigma_l \leqq m-\sigma_k} \frac{f_{k+l}(x)}{l!}(x'-x)^l + R_k(x';x).$$

Let x^0 be an arbitrary point of E. If for each k ($\sigma_k \leqq m$) and every $\epsilon>0$ there is a $\delta>0$ such that

$$(2) \qquad |R_k(x';x)| \leqq r_{xx'}^{m-\sigma_k}\epsilon \text{ if } x, x' \text{ in } A,\ r_{xx^0}<\delta,\ r_{x'x^0}<\delta,$$

we shall say that $f(x)$ is *of class C^m in A about x^0 in terms of the $f_k(x)$*, or, $f(x)$ is $(C^m, A, x^0, f_k(x))$. If this is true for each x^0 in B, we say $f(x)$ is $(C^m, A, B, f_k(x))$, and replace "about x^0" by "about B." We say $f(x)$ (defined in A) is *of class C^m in A about B*, or, $f(x)$ is (C^m, A, B), if there exist functions $f_k(x)$ ($\sigma_k \leqq m$) defined in A such that $f(x)$ is $(C^m, A, B, f_k(x))$. If $B=A$ in the last two definitions, we leave out the words "about B"; this is in agreement with the previous definitions. We say $f(x)$ is $(C^\infty, A, B, f_k(x))$ if $f(x)$ is $(C^m, A, B, f_k(x))$ for each m. Any function defined in A is (C^{-1}, A, E).

Remark. We might define in an obvious manner such relations as (C^m, A, x^0), (C^∞, A, B). To study them would require a study of the different possible definitions of the $f_k(x)$ if $f(x)$ is (C^m, A, B). The $f_k(x)$ are not in general determined by $f(x)$. Thus if $A=B$ is the x_1-axis, only the $f_k(x)$ with $k_2=\cdots=k_n=0$ are determined by $f(x)$. It is not obvious for what point sets A the $f_k(x)$ are all determined by $f(x)$.

If $f(x)$ is $(C^m, A, B, f_k(x))$ ($m \geqq 0$), then the $f_k(x)$ are continuous at each point of B;* that is, the $f_k(x)$ may be defined in $B-B\cdot A$ so that this will be true. To show this, take x^0 in B, set $\epsilon=1$, and choose δ so that (2) holds for any k ($\sigma_k \leqq m$). Take x in A within δ of x^0 (if there is such a point); then (1) and (2) show that $f_k(x')$ is bounded for x' in A within δ of x^0 ($\sigma_k \leqq m$). Now let $\{x^i\}$ be any sequence of points of A, $x^i \to x^0$; (1) and (2) show that $\{f_k(x^i)\}$ is a regular sequence.

If A is open and $f(x)$ is $(C^m, A, A, f_k(x))$, then $D_k f(x)$ exists and equals $f_k(x)$ in A ($\sigma_k \leqq m$). (See AE.) If x^0 is an isolated point of A or is at a positive distance from A, then $f(x)$ is $(C^m, A, x^0, f_k(x))$ for any $f_k(x)$. If $f(x)$ is $(C^m, A, B, f_k(x))$ [or (C^m, A, B)], and A' is in A, B' is in B, then $f(x)$ is $(C^m, A', B', f_k(x))$ [or (C^m, A', B')]. Also $f(x)$ is (C^0, A, B) if and only if it is continuous at each point of B. If $f(x)$ is $(C^m, A, B, f_k(x))$, then it is $(C^{m'}, A, B, f_k(x))$ for all $m'<m$; a stronger theorem is proved in Theorem 2. If $f(x)$ is $(C^m, A, B, f_k(x))$, then $f_k(x)$ is $(C^{m-\sigma_k}, A, B, f_l(x))$.

3. **Extension theorems.** We prove here a theorem which gives the maximum range of differentiability of a function, and a theorem about the still larger range of differentiability of a function to an order one less.

* Or better, "continuous in A about B."

THEOREM 1. *If $f(x)$ is $(C^m, A, B, f_k(x))$* (m finite or infinite), then the $f_k(x)$ may be extended throughout E so that $f(x)$ is $(C^m, E, B, f_k(x))$.*†

We note, conversely, that if $f(x)$ is not $(C^m, A, x^0, f_k(x))$, then no extension of $f(x)$ will be so. We remark also that $f(x)$ may be made analytic in $E-\overline{A}$ ($\overline{A}=A$ plus limit points).

To prove the theorem, we first extend the $f_k(x)$ through $\overline{A}-A$ as follows: Take any x^0 in $\overline{A}-A$. Let $f_k(x^0)$ be the upper limit of $f_k(x^i)$ for sequences $\{x^i\}$, $x^i \to x^0$, x^i in A, if this is finite; otherwise, set $f_k(x^0)=0$. Next we extend the $f_k(x)$ throughout $E-\overline{A}$ by the method of AE Lemma 2. We shall assume in the proof that m is finite. If $m=\infty$, we prove $C^{m'}$ for every integer m'. The only alteration needed in the proof is that AE §12 should be used; but this makes no essential change.

As $E-\overline{A}$ is open, $f(x)$ is $(C^m, E, E-\overline{A}, f_k(x))$; we must show that $f(x)$ is $(C^m, E, B\cdot\overline{A}, f_k(x))$. Take a fixed point x^0 in $B\cdot\overline{A}$. Let us say (k, ϵ, A_1, A_2) holds if there is a $\delta>0$ such that (2) holds whenever x is in A_1, x' is in A_2, and $r_{xx^0}<\delta$, $r_{x'x^0}<\delta$. We must prove (k, ϵ, E, E) for each k ($\sigma_k \leqq m$) and each $\epsilon>0$.

First we prove $(k, \epsilon, \overline{A}, \overline{A})$. Set $\epsilon'=\epsilon/[2(m+1)^n]$, and let δ be the smallest of the δ's given by (l, ϵ', A, A) for $\sigma_l \leqq m$. Let U be the spherical neighborhood of x^0 of radius δ; then $f_l(x)$ is bounded in $U\cdot A$ ($\sigma_l \leqq m$). Given x, x' in $U\cdot\overline{A}$, choose sequences $\{x^i\}$, $\{x'^i\}$ of points of $U\cdot A$, with $x^i \to x$, $x'^i \to x'$. Suppose first $\sigma_k=m$. Then we may take these sequences so that $f_k(x^i)\to f_k(x)$, $f_k(x'^i)\to f_k(x')$, and the desired inequality for $R_k(x'; x)$ follows from that for $R_k(x'^i; x^i)$. Suppose now that $\sigma_k<m$. Relations (1) and (2) with k, x', x replaced by l, x^i, x^j show that for any such $\{x^i\}$, $\{f_l(x^i)\}$ is a regular sequence ($\sigma_l<m$); hence $f_l(x^i)\to f_l(x)$, and similarly $f_l(x'^i)\to f_l(x')$ ($\sigma_l<m$). Relation (1) now shows that for i large enough, $\Delta=R_k(x'; x)-R_k(x'^i; x^i)$ differs as little as we please from

$$-\sum_{\sigma_l=m-\sigma_k} \frac{f_{k+l}(x)-f_{k+l}(x^i)}{l!}(x'-x)^l.$$

As $|f_j(x)-f_j(x^i)| \leqq \epsilon'$ ($\sigma_j=m$) and $|(x'-x)^l| \leqq r_{xx'}^{\sigma_l}$, $|\Delta| \leqq (m+1)^n\epsilon' r_{xx'}^{m-\sigma_k}$ for i large enough; the inequality again follows.

Next we prove $(k, \epsilon, \overline{A}, E-\overline{A})$. Set $\epsilon'=\epsilon/[2\cdot 4^m(m+1)^n]$, and define η in terms of ϵ' and then δ as in AE §11, using $(k, \eta, \overline{A}, \overline{A})$. Take x in $\overline{A}$ and x' in $E-\overline{A}$, each within $\delta/4$ of x^0. By AE (6.3) and the equation following (11.6),

* Or merely locally (C^m, A, B); see Theorem 6.

† If $A=B$ is closed, then B may be replaced by E; the present proof then gives a proof of AE Lemma 2 which makes no use of AE Lemma 1.

$$R_k(x';\, x) = D_k f(x') - \psi_k(x';\, x)$$
$$= \sum_l \frac{R_{k+l}(x^*;\, x)}{l!}(x' - x^*)^l + \sum_{s=1}^{t}\sum_l \binom{k}{l} D_l\phi_{\lambda_s}(x')\zeta_{\lambda_s;k-l}(x'),$$

where x^* is a point of $\overline{A}$ distant $\delta_*/4$ from x', $\delta_*/4$ being the distance from x' to $\overline{A}$. As $r_{x^*x} \leqq 2r_{xx'}$, $r_{x'x^*} \leqq 2r_{xx'}$, and $\delta_* \leqq 4r_{xx'}$, we find with the help of AE (11.8)

$$|\, R_k(x';\, x)\,| \leqq (m+1)^n (2r_{xx'})^{m-\sigma_k}\eta + (4r_{xx'})^{m-\sigma_k}\epsilon'/2 < r_{xx'}^{m-\sigma_k}\epsilon.$$

Next we prove $(k,\ \epsilon,\ E-\overline{A},\ \overline{A})$. As is easily seen from AE (6.3) or by F (6) with x^{i-1}, x^i replaced by x, x',

$$R_k(x';\, x) = \sum_l \frac{R_{k+l}(x;\, x')}{l!}(x' - x)^l.$$

Set $\epsilon' = \epsilon/(m+1)^n$, and take the smallest δ given by $(k+l,\ \epsilon',\ \overline{A},\ E-\overline{A})$ for $\sigma_l \leqq m-\sigma_k$. The required inequality now follows at once.

Finally we must show $(k,\ \epsilon,\ E-\overline{A},\ E-\overline{A})$. Set $\epsilon' = \epsilon/[2n(m+1)^n]$, and take δ smaller than the $\delta/4$ given by AE §11 with ϵ replaced by ϵ' and smaller than the δ's given by $(k+l,\ \epsilon',\ \overline{A},\ E-\overline{A})$ and $(k+l,\ \epsilon',\ E-\overline{A},\ \overline{A})$ for $\sigma_l \leqq m-\sigma_k$. Now take x and x' in $E-\overline{A}$ within δ of x^0; we must consider two cases. Case I: The line segment $S = xx'$ lies wholly in $E-\overline{A}$. By AE (11.2), $|f_l(y) - f_l(x')| < 2\epsilon'$ for y on S $(\sigma_l \leqq m)$; the desired inequality now follows from F, Lemma 3. Case II: There is a point x^* of $\overline{A}$ on S. From AE (6.3), or F (6) with x^{i-1}, x^i replaced by x, x^*, we find

$$R_k(x';\, x) = R_k(x';\, x^*) + \sum_l \frac{R_{k+l}(x^*;\, x)}{l!}(x' - x^*)^l,$$

and the inequality again follows.

THEOREM 2. *If $f(x)$ is $(C^m,\ A,\ B,\ f_k(x))$ (m finite), then there is an open set B' containing B such that $f(x)$ is $(C^{m-1},\ A,\ B',\ f_k(x))$.*

For each x in B, let $\delta(x)$ be the largest of the numbers δ for which (2) holds for all k $(\sigma_k \leqq m)$ with ϵ replaced by 1. Let $U(x)$ be the set of all points x' within $\delta(x)$ of x; then B' is the sum of all $U(x)$. The set B' is open. To prove $(C^{m-1},\ A,\ B',\ f_k(x))$, take any x^0 in B' and any $\epsilon > 0$. For some x^* in B, $r_{x^*x^0} < \delta(x^*)$. There is an M such that $|f_k(y)| < M$ for y in $A \cdot U(x^*)$ $(\sigma_k \leqq m)$.† Let δ be the smaller of $\delta(x^*) - r_{x^*x^0}$ and $\epsilon/[2(m+1)^n M + 2]$. Now take any x and x' in A within δ of x^0. We are interested in the remainders

† For the proof, see the paragraph following the remark.

$$R_k'(x'; x) = \sum_{\sigma_l = m - \sigma_k} \frac{f_{k+l}(x)}{l!} (x' - x)^l + R_k(x'; x)$$

with $\sigma_k < m$. As $r_{xx'} < 2\delta$,

$$| R_k'(x'; x) | \leqq (m+1)^n M r_{xx'}^{m-\sigma_k} + r_{xx'}^{m-\sigma_k} < r_{xx'}^{m-1-\sigma_k} \epsilon.$$

COROLLARY. *If $f(x)$ is of class C^m in any given point set about B, then it may be extended through an open set B' containing B so that it is of class C^{m-1} in B' and of class C^m in B' about B.*

4. **Composite functions, etc.** We prove here three theorems.

THEOREM 3. *If f and g are of class C^m in A about B, then so are $f+g$ and $f-g$, with*

$$(f \pm g)_k = f_k \pm g_k. \tag{3}$$

This is obvious.

THEOREM 4. *If f and g are of class C^m in A about B, then so is fg, and f/g if $g \neq 0$. The derivatives are given by the ordinary formulas. Thus*

$$(fg)_k = \sum_l \binom{k}{l} f_l g_{k-l}. \tag{4}$$

We might prove this theorem directly, but it follows from Theorem 5: fg and f/g are functions (of two variables) of class C^∞ of the functions f and g. (The condition B in A is obtained by using Theorem 1.)

THEOREM 5. *Let A and B be subsets of n-space E_n, and let A' and B' be subsets of ν-space E_ν. Let $f^i(x)$ be $(C^m, A, B, f_k^i(x))$ $(i = 1, \cdots, \nu)$, and let $g(y)$ be $(C^m, A', B', g_k(y))$ (m finite or infinite). Suppose B is in A, x in A implies*

$$y = (y_1, \cdots, y_\nu) = (f^1(x), \cdots, f^\nu(x)) = f(x)$$

in A', and x in B implies $f(x)$ in B'. Then the function

$$h(x) = g(f^1(x), \cdots, f^\nu(x)) = g(f(x))$$

is $(C^m, A, B, h_k(x))$; the $h_k(x)$ are given by the ordinary formulas (9) for derivatives.

As a consequence of this theorem, the definition of being of class C^m is independent of the coordinate system chosen. If the condition x in A [or B] does not imply $f(x)$ in A' [or B'], we may apply the theorem to any subset A_1 [or B_1] of A [or B] for which it does. We shall suppose m is finite; if $m = \infty$, we merely apply the reasoning below for each positive integer.

Suppose first $u^1(x), \cdots, u^r(x)$ are functions of class C^m in an open set Γ of E_n, suppose $v(y)$ is of class C^m in an open set Γ' of E_r, and suppose x in Γ implies $u(x)$ in Γ'. Letting R'^i, S' denote remainders for u^i, v, Taylor's formula gives

$$u_k^i(x') = D_k u^i(x') = \sum_{\sigma_l \leqq m-\sigma_k} \frac{u_{k+l}^i(x)}{l!}(x'-x)^l + R_k'^i(x';x), \tag{5}$$

$$v_k(y') = D_k v(y') = \sum_{\sigma_l' \leqq m-\sigma_k'} \frac{v_{k+l}(y)}{l!}(y'-y)^l + S_k'(y';y), \tag{6}$$

certain inequalities on the $R_k'^i$ and S_k' being satisfied. We have set $\sigma_k' = k_1 + \cdots + k_r$. Set $w(x) = v(u(x))$; then (5) and (6) with $k=0$ give

$$w(x') = \sum_t \frac{v_t(u(x))}{t!}\left\{\sum_{\sigma_j \geqq 1} \frac{u_j(x)}{j!}(x'-x)^j + R'(x';x)\right\}^t + S'(u(x');u(x)), \tag{7}$$

where $S' = S_0'$. Also, by Taylor's formula,

$$w_k(x') = \sum_l \frac{w_{k+l}(x)}{l!}(x'-x)^l + T_k'(x';x). \tag{8}$$

Subtract (8) with $k=0$ from (7); then as R'^i, S', and T' all approach 0 to the mth order as $x' \to x$,† we may equate coefficients of $(x'-x)^k$ for $\sigma_k \leqq m$.‡ Thus we find polynomials

$$P_k(u_p^i, v_q) \qquad (\sigma_p \leqq \sigma_k,\ \sigma_q' \leqq \sigma_k;\ \sigma_k \leqq m)$$

such that, for any x in Γ,

$$w_k(x) = P_k(u_p^i(x), v_q(u(x))). \tag{9}$$

Using (8) gives for $w_k(x')$

$$w_k(x') = \sum_l \frac{P_{k+l}(u_p^i(x), v_q(u(x)))}{l!}(x'-x)^l + T_k'(x';x). \tag{10}$$

We may also evaluate it by replacing x by x' in (9) and using (5) and (6). (In (6) we replace y' by $u(x')$ and use (5) again.) Each variable in the resulting polynomial P_k consists of a polynomial in quantities R', S', and other quantities; if we multiply out and collect all terms with an R' or an S' as a factor, we obtain

† This is clear for S' if $m=0$; if $m>0$, then

$$S'/r_{xx'}^m = [S'/|u(x')-u(x)|^m]\cdot[|u(x')-u(x)|/r_{xx'}]^m,$$

where $|y'-y| = r_{yy'}$, and the last factor is bounded in $U \cdot A$.

‡ This is easily proved in succession for $\sigma_k = 0, 1, \cdots$ on letting $x' \to x$.

$$w_k(x') = P_k\Big[\sum_s \frac{u_p^{i}{}_{+s}(x)}{s!}(x'-x)^s, \sum_t \frac{v_{q+t}(u(x))}{t!}\Big\{\sum_{\sigma_j\geqq 1}\frac{u_j(x)}{j!}(x'-x)^j\Big\}^t\Big] + Q_k, \tag{11}$$

where Q_k is a polynomial containing an R' or an S' as a factor in each term. It must be understood that $\sum u_p^{i}{}_{+s}(x)(x'-x)^s/s!$ appears as the variable in the position of u_p^i, etc., in $P_k(u_p^i, v_q)$.

We now prove: *If u_k^i $(\sigma_k \leqq m; i=1, \cdots, \nu)$, v_k $(\sigma_k' \leqq m)$ are any numbers, then*

$$P_k^*(x; u_p^i, v_q) = P_k\Big[\sum_s \frac{u_p^{i}{}_{+s}}{s!}x^s, \sum_t \frac{v_{q+t}}{t!}\Big\{\sum_{\sigma_j\geqq 1}\frac{u_j}{j!}x^j\Big\}^t\Big] - \sum_l \frac{P_{k+l}(u_p^i, v_q)}{l!}x^l, \tag{12}$$

considered as a polynomial in x, contains no terms of degree $\leqq m-\sigma_k$. To prove this, define the polynomials

$$u^i(x) = \sum_{\sigma_l\leqq m}\frac{u_l^i}{l!}x^l, \qquad v(y) = \sum_{\sigma_l'\leqq m}\frac{v_l}{l!}(y-u_0)^l; \tag{13}$$

then $u_k^i(0) = D_k u^i(0) = u_k^i$, $v_k(u_0) = D_k v(u_0) = v_k$. Set $w(x) = v(u(x))$. Replacing x', x by x, 0 in (10) and (11) and putting in (12) gives, as $Q_k = 0$ in this case,

$$P_k^*(x; u_p^i, v_q) = T_k'(x; 0). \tag{14}$$

As $T_k' \to 0$ to the $(m-\sigma_k)$th order as $x \to 0$, P_k^* cannot contain any terms of degree $\leqq m-\sigma_k$.

We return now to the functions $f^i(x)$, $g(y)$, $h(x)$. Set $h_k(x) = P_k(f_p^i(x), g_q(f(x)))$. The formulas (10) and (11) hold equally well for the f^i, g, h. Hence using (10), (11), and (12), we find for the remainder for $h_k(x)$

$$T_k(x'; x) = P_k^*(x' - x; f_p^i(x), g_q(f(x))) + Q_k. \tag{15}$$

To show that $h(x)$ is $(C^m, A, B, h_k(x))$, take any x^0 in B, and set $y^0 = f(x^0)$. As $f(x)$ is continuous in A about B, for each neighborhood V of y^0 there is a neighborhood $U(V)$ of x^0 such that x in $U(V)\cdot A$ implies $f(x)$ in $V\cdot A'$. As y^0 is in B', we may take V so that the $g_k(y)$ are bounded in $V\cdot A'$. We may take U in $U(V)$ so small that the $f_k(x)$ are bounded in $U\cdot A$. Because of the property of P_k^*, we may obviously take δ small enough so that P_k^* satisfies an inequality of the nature of (2). Moreover each term in Q_k contains an $R_p(x'; x)$ or an $S_q(u(x'); u(x))$ with $\sigma_p \leqq \sigma_k$ or $\sigma_q' \leqq \sigma_k$; as each such remainder satisfies

an inequality (2) (see a recent footnote) and all other quantities entering into Q_k are bounded, we may take δ small enough so that Q_k also satisfies an inequality (2). Hence the same is true of T_k, and the theorem is proved.

5. **Differentiability a local property.** Our object is to prove

THEOREM 6. *Let $f(x)$ be locally (C^m, A, B) (m finite or infinite). For each point x^0 of B there is a neighborhood U of x^0 and functions $f_k^{(x^0)}(x)$ defined in $U \cdot A$ such that $f(x)$ is $(C^m, U \cdot A, U \cdot B, f_k^{(x^0)}(x))$.† Then $f(x)$ is (C^m, A, B). If the $f_k^{(x^0)}(x)$ for $\sigma_k \leqq p$ are independent (at any x for which they are defined) of x^0, then these functions may be included among the $f_k(x)$ $(\sigma_k \leqq m)$.*

We may take each neighborhood U as an open n-cube, so small that the $f_k^{(x^0)}(x)$ are bounded in U. A finite or denumerable number of them, $C_1, C_2, \cdots$, cover B; we may take them so that any one touches at most a finite number of the others, and so that any boundary point of any C_i is interior to some C_j.‡ By hypothesis, to each i there correspond functions $f_k^i(x)$, $\sigma_k \leqq m$, such that $f(x)$ is $(C^m, C_i \cdot A, C_i \cdot B, f_k^i(x))$. In each C_i we define the function $\pi_i(x)$ as it was defined in I_i in AE §9; set

$$\phi_i(x) = \pi_i(x) / \sum_j \pi_j(x) \tag{16}$$

in $C_1 + C_2 + \cdots$. Set $g^i(x) = \phi_i(x) f(x)$ in $C_i \cdot A$. By Theorem 4, $g^i(x)$ is $(C^m, C_i \cdot A, C_i \cdot B)$, and

$$g_k^i(x) = \sum_l \binom{k}{l} D_l \phi_i(x) f_{k-l}^i(x). \tag{17}$$

As the $f_k^i(x)$ are bounded in $C_i \cdot A$ and the $D_l \phi_i(x) \to 0$ to infinite order as x approaches the boundary of C_i (see AE §9), the latter statement is true also of the $g_k^i(x)$. Hence, evidently, if we set $g_k^i(x) = 0$ in $A - C_i \cdot A$, $g^i(x)$ is $(C^m, A, B, g_k^i(x))$. Set

$$f_k(x) = g_k^1(x) + g_k^2(x) + \cdots, \tag{18}$$

which in any $C_i \cdot A$ is a finite sum; this reduces to $f(x)$ for $k = 0$. Theorem 3 shows at once that $f(x)$ is $(C^m, A, B, f_k(x))$. (Given x^0 in B, to apply Theorem

† Note that "$f(x)$ is locally $(C^\infty, \cdots)$" is not the same statement as "$f(x)$ is locally $(C^m, \cdots)$ for each m."

‡ Let $C^1, C^2, \cdots$ be a denumerable set of the cubes which cover B. Express each C^i as the sum of a denumerable number of cubes C_j^i with the following properties: Each C_j^i is, with its boundary, interior to C^i; the diameter of C_j^i, $\delta(C_j^i)$, is $< 1/i$; $\delta(C_j^i) \to 0$ as $j \to \infty$; the cubes C_j^i approach the boundary of C^i as $j \to \infty$. Now drop out all cubes C_j^i which are interior to larger cubes C_l^k; the remaining cubes $C_1, C_2, \cdots$ still cover B. To each cube C_j^i corresponds a number $\eta > 0$ such that any point set of diameter $< \eta$ having points in common with C_j^i lies interior to some C_k^i; using this fact, it is easily seen that any C_i has points in common with but a finite number of the C_j.

3, we choose δ so small that the points within δ of x^0 lie in but a finite number of the C_i.)

To prove the second statement, let $f_l'(x)$ denote the common value of $f_l^i(x)$ for $\sigma_l \leqq p$. Differentiating $\sum \phi_i = 1$ gives

$$\sum_i D_l\phi_i(x) = \begin{cases} 1 \text{ if } l = 0, \\ 0 \text{ if } \sigma_l > 0. \end{cases} \tag{19}$$

Define the $f_k(x)$ as before. Take any k with $\sigma_k \leqq p$; then (17) and (18) give

$$f_k(x) = \sum_l \binom{k}{l} f_{k-l}'(x) \sum_i D_l\phi_i(x) = f_k'(x)$$

in $C_1+C_2+\cdots$. It does not matter how $f_k(x)$ is defined outside this set.

The second statement in the theorem does not hold for an arbitrary set of $f_k(x)$, at least using the above method. To see this, take $n=m=2$, $A=B$ =the interval $(-1, 1)$ of the x_1-axis, $C_1=C_2$=the square with corners $(\pm 1, \pm 1)$; set $f=0$,

$$f_{10}^1 = f_{20}^1 = f_{11}^1 = f_{02}^1 = 0, \qquad f_{01}^1 = 1,$$

and $f_{ij}^2 = -f_{ij}^1$ on A. Also set

$$\phi_1(x, y) = \tfrac{1}{2} + \tfrac{3}{4}x - \tfrac{1}{4}x^3, \qquad \phi_2(x, y) = \tfrac{1}{2} - \tfrac{3}{4}x + \tfrac{1}{4}x^3.$$

(Though ϕ_1 and ϕ_2 are not the functions defined above, they have the necessary properties.) We find on A

$$g_{11}^1(x, y) = g_{11}^2(x, y) = \tfrac{3}{4} - \tfrac{3}{4}x^2, \qquad f_{11}(x, y) = \tfrac{3}{2} - \tfrac{3}{2}x^2 \not\equiv 0.$$

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.

ON THE CLASS NUMBER OF A QUATERNION ALGEBRA WITH A NEGATIVE FUNDAMENTAL NUMBER*

BY

CLAIBORNE G. LATIMER

1. **Introduction.** Let $\mathfrak{A}$ be a rational generalized quaternion algebra and let $\mathfrak{G}$ be a set of integral elements in $\mathfrak{A}$ according to Dickson's definition.† The number h, of classes of left ideals in $\mathfrak{G}$ is independent of the particular $\mathfrak{G}$ in $\mathfrak{A}$ which is considered and is equal to the number of classes of right ideals.‡ We shall show that if $\mathfrak{A}$ contains an element with a negative norm, then every ideal in $\mathfrak{G}$ is principal and hence $h=1$.

From this, we obtain theorems on the existence of a greatest common right divisor, or g.c.r.d., and a g.c.l.d. of two elements and on the factorization of an element in $\mathfrak{G}$. Results equivalent to these have been previously obtained for a number of special cases.§ These theorems are also similar to well known results, due to Hurwitz, for a certain $\mathfrak{G}$ in the classic rational quaternion algebra.

Let G be the set of all integral algebraic numbers in a quadratic field. It will be shown that G is contained in a set $\mathfrak{G}$ as above defined, such that every pair of elements ξ, η in G, not both zero, have a g.c.r.d. δ_1 and a g.c.l.d. δ_2 in $\mathfrak{G}$ which are uniquely determined, apart from unit factors. Moreover, if ξ, η have a g.c.d. δ in G, we may take $\delta_1=\delta_2=\delta$. Hence, while sacrificing commutativity, such an enlargement of G yields a result similar to that obtained by the introduction of ideals.

2. **A special $\mathfrak{G}$.** It is known that $\mathfrak{A}$ has a basis $1, i, j, ij$ such that

$$i^2 = -\tau, \qquad j^2 = -d, \qquad ij = -ji,$$

where d is the fundamental number of $\mathfrak{A}$ and τ is any one of the infinitude of positive primes which satisfy the conditions:

(a) τ is prime to d, $\tau \equiv 3 \pmod 4$ and $-\tau$ is a quadratic non-residue of every odd prime factor of d;

* Presented to the Society, September 13, 1935, received by the editors October 9, 1935.

† *Algebras and their Arithmetics*, pp. 141–142. This definition is equivalent to the definition of a maximal realm of integrity as given by Brandt, *Idealtheorie in Quaternionenalgebren*, Mathematische Annalen, vol. 99 (1928), p. 11.

‡ Brandt, loc. cit., p. 23.

§ E.g., *On ideals in generalized quaternion algebras and Hermitian forms*, these Transactions, vol. 38 (1935), pp. 436–446, where references are given to other special cases. This paper will be referred to hereafter as Tr.

(b) $-d$ is a quadratic residue of τ;

(c) if d is even, $\tau \equiv 3 \pmod 8$.*

Though not necessary for the proofs of the three lemmas to follow, we shall assume without loss of generality that

$$\tau > 4|d|. \tag{d}$$

We shall now prove

LEMMA 1. *Let μ be an integer such that $4\mu^2+d\equiv 0 \pmod\tau$. Then*

$$e_1 = 1, \quad e_2 = \frac{\tau + i}{2}, \quad e_3 = j, \quad e_4 = \frac{(-2\mu + e_3)e_2}{\tau}$$

form a basis of a set of integral elements.†

We shall prove this lemma by showing that the e's may be obtained by an integral transformation of determinant ± 1 from known basal elements of a set of integral elements.

For the case where d is odd, basal elements of those sets of integral elements which contain i and j may be obtained by specializing certain results obtained by the writer.‡ In this reference, set $I=i, J=j, K=ij, \alpha=-\tau, \beta=d$, take H_1 so that $dH_1\equiv 2\mu \pmod\tau$ and, employing (a) and (b) above, properly specialize the various divisors of $\alpha\beta$ which appear in the expressions for the basal elements 1, P'', Q_1'', R_1 of S_1'', as given on p. 62. It will then be found that the above e's are obtained from these basal elements by an integral transformation of determinant ± 1. But by the same reference, S_1'' is a set of integral elements. The lemma follows for the case where d is odd.

Suppose d is even. We employ results due to Darkow,§ setting her e_1, e_2, e_3 equal to $ij, j, -di$, respectively. Then by (c) above, $\mathfrak{A}$ is of type A as defined by Darkow. Take her $H_1=F_1=1$ and E_1 so that $dE_1\equiv 2\mu \pmod\tau$. It will be found, as in the case where d is odd, that the e's of the lemma are obtained from her $A_0, \cdots, A_3$ (p. 267) by an integral transformation of determinant ± 1. But the A's form a basis of a set of integral elements. The lemma follows.

---

* *On the fundamental number of a rational generalized quaternion algebra*, Duke Mathematical Journal, vol. 1 (1935), pp. 433–435. For the definition of d, see Brandt, loc. cit., p. 9.

† For the case where $\mathfrak{A}$ is a division algebra, a similar result was obtained by Albert. See his *Integral domains of rational generalized quaternion algebras*, Bulletin of the American Mathematical Society, vol. 40 (1934), p. 176, Theorem 9. In the paper in the Duke Mathematical Journal cited above, it was shown that Albert's $\sigma=-d$. Albert's τ, the negative of ours, is subject to the restriction that it is represented by a certain binary quadratic form.

‡ *Arithmetics of generalized quaternion algebras*, American Journal of Mathematics, vol. 48 (1926), pp. 57–63.

§ *Determination of a basis for the integral elements of certain generalized quaternion algebras*, Annals of Mathematics, (2), vol. 28 (1926–27), pp. 263–270.

It will be assumed hereafter, unless the contrary is explicitly stated, that $\mathfrak{G}$ is the set of integral elements given by Lemma 1.

Suppose $\mathfrak{A}$ has a basis $1, I, J, IJ$, where $I^2=-\alpha$, $J^2=-\beta$, $IJ=-JI$, α and β being rational integers, neither divisible by the square of a prime. It may be shown that if a set of integral elements contains two of $I, J, IJ/\delta$, then it contains the third, δ being the g.c.d. of α and β, for otherwise the set would not be maximal. While not necessary for the sequel, we note that this, together with Lemma 1 and the papers cited in its proof, has an important bearing on certain remarks and results in Albert's paper previously cited.

In the papers cited in the proof of Lemma 1, Darkow and the writer determined those sets of integral elements containing I and J if $\alpha\equiv\beta \pmod 2$, Darkow treating the case where α is even. Albert remarked (loc. cit., p. 165) that these results did not complete the problem, presumably referring to the case where $\alpha\equiv\beta+1 \pmod 2$. However this lack of completeness is only apparent since, by a change of notation, this case is seen to be included in Darkow's results. Thus if $\alpha\equiv\beta+1\equiv 1 \pmod 2$, the sets containing I and J are identical with the sets containing IJ/δ and J, and the latter are obtained from Darkow's results by setting her e_1, e_2, e_3 equal to IJ/δ, J, $-\beta I/\delta$, respectively.

If d is odd, by the above mentioned results of the writer's, there are exactly two sets of integral elements containing e_2 and e_3 of Lemma 1. If d is even, the same is true by Darkow's results and by the above statement as to a set containing two of $I, J, IJ/\delta$. Albert's Theorem 9 (loc. cit., p. 176) follows at once from Lemma 1 without the restriction that $\mathfrak{A}$ be a division algebra.

3. **On the ideals in a certain ring.** It will be observed that 1 and e_2 form a basis of a set G, equivalent to the set of all integral numbers in the field defined by $(-\tau)^{1/2}$. Let $\mathfrak{G}_1$ be the ring consisting of all elements in the form $x+e_3y$, where x, y are in G. $\mathfrak{G}$ contains $\mathfrak{G}_1$ and an element $\sum u_ie_i$ of $\mathfrak{G}$ is in $\mathfrak{G}_1$ if and only if $u_4\equiv 0 \pmod\tau$. Hence $\bar{e}_i=e_i$ $(i=1, 2, 3)$, $\bar{e}_4=\tau e_4$ form a basis of $\mathfrak{G}_1$.

We shall use the same definitions of a left ideal in $\mathfrak{G}$, a class of such ideals and the norm of an ideal as given by Brandt.* A left ideal in $\mathfrak{G}_1$, a regular ideal and a class of such ideals are defined as in Tr.† Unless the contrary is explicitly stated, it will be understood that all ideals referred to, in $\mathfrak{G}$ or in $\mathfrak{G}_1$, are left ideals; also, that every element of an ideal in $\mathfrak{G}$ belongs to $\mathfrak{G}$.

Those elements of an ideal $\mathfrak{L}$, in $\mathfrak{G}$ which are in $\mathfrak{G}_1$ form an ideal $\mathfrak{L}_1$, in $\mathfrak{G}_1$ which will be said to correspond to $\mathfrak{L}$. We shall now prove

* Brandt, loc. cit., pp. 16, 14. See also second footnote, p. 21 and p. 3.

† To identify the notation of Tr. with that of this paper, set $E=e_3, \alpha=-d, \Delta=-\tau, \mathfrak{G}=\mathfrak{G}_1$.

LEMMA 2. *If* $\mathfrak{L}$ *is an ideal in* $\mathfrak{G}$ *and* $\mathfrak{L}_1$ *is the corresponding ideal in* $\mathfrak{G}_1$, *then* $\mathfrak{L}_1$ *is a regular ideal if and only if* $\mathfrak{L}$ *contains an element not in* $\mathfrak{G}_1$.

$\mathfrak{L}$ has a basis $\omega_i=\sum_j a_{ij}e_j$ $(i=1, \cdots, 4)$, where the determinant $|a_{ij}|=a^2$, a being a positive integer which, by definition, is the norm $N(\mathfrak{L})$ of $\mathfrak{L}$.* We shall assume, without loss of generality, that every $a_{ii}>0$ and $a_{ij}=0$ if $i<j$.

$\mathfrak{L}_1$ has as a basis $\bar{\omega}_i=\omega_i$ $(i=1, 2, 3)$, $\bar{\omega}_4=\tau\omega_4$ or $\bar{\omega}_i=\omega_i$ $(i=1, \cdots, 4)$ according as not all or all of the elements of $\mathfrak{L}$ are in $\mathfrak{G}_1$. If we set $\bar{\omega}_i=\sum_j \bar{a}_{ij}\bar{e}_j$, in the first case the determinant $|a_{ij}|=|\bar{a}_{ij}|$ and in the second case, $|a_{ij}|=\tau|\bar{a}_{ij}|$.

There are certain ideals $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{d}$ in G determined by $\mathfrak{L}_1$ and by Lemmas 1, 2 of Tr., $|\bar{a}_{ij}|=N(\mathfrak{a})N(\mathfrak{b})=a_1^2N^2(\mathfrak{b})N(\mathfrak{d})$, where a_1 is a rational integer. $-\tau$ is the discriminant of the quadratic field defined by i. Hence by (a) and (c) of §2 and by a well known result,† no rational prime factor of d is divisible by a prime ideal in G of the first degree. Then by Lemma 2 of Tr., $N(\mathfrak{d})=1$ or $N(\mathfrak{d})=\tau$. But $|a_{ij}|$ is a square. Therefore in the first case above $N(\mathfrak{d})=1$; in the second case, $N(\mathfrak{d})=\tau$. Since $\mathfrak{L}_1$ is regular by definition if and only if $N(\mathfrak{d})=1$, the lemma follows.

We shall assume hereafter that $d<0$.

LEMMA 3. *Every regular ideal in* $\mathfrak{G}_1$ *is a principal ideal.*

By Theorem 3 of Tr. there is a one-to-one correspondence between the regular classes of ideals in $\mathfrak{G}_1$ and the classes of Hermitian forms in G of determinant $-d$. By the last sentence in §4 of Tr., every ideal in a regular class is regular. Hence, to prove the lemma, it is sufficient to show that there is a single class of such Hermitian forms. Let $f(x, y)=axx'+bx'y+b'xy'+cyy'$ be such a form. It may be shown that f is equivalent, under a unitary transformation with coefficients in G, to such a form with a odd and b in the ring G_1, consisting of all elements in the form $r+si$ where r, s are rational integers. Assume that a is odd and b is in G_1. Since $bb'-ac=-d$ and d contains no square factor >1,‡ if we restrict x, y to G_1, f is a properly primitive form in G_1. By a result due to Humbert,§ there is a single class of such forms; i.e., any two such forms are equivalent under a unitary transformation with coefficients in G_1. Since G contains G_1, the lemma follows.

4. **Proof of principal theorem.** Let C be a class of ideals in $\mathfrak{G}$. The class

* Brandt, loc. cit., pp. 13, 16.

† Cf. Hecke, *Vorlesungen über die Theorie der Algebraischen Zahlen*, p. 110, Satz. 90.

‡ Fueter, *Zur Theorie der Brandtschen Quaternionenalgebren*, Mathematische Annalen, vol. 110 (1934), p. 658.

§ Comptes Rendus, Paris, vol. 171 (1920), pp. 287–293; Dickson, *History of the Theory of Numbers*, vol. 3, p. 278.

C contains an ideal $\mathfrak{L}$ such that $N(\mathfrak{L}) \leqq 2(-d)^{1/2}$.* Let $\omega_1, \cdots, \omega_4$ be a basis of $\mathfrak{L}$ as used in the proof of Lemma 2. Then by (d) of §2, $N^2(\mathfrak{L}) = a_{11}a_{22}a_{33}a_{44} \leqq -4d < \tau$, a_{44} is prime to τ and $\bar{\omega}_i = \omega_i$ $(i = 1, 2, 3)$, $\bar{\omega}_4 = \tau\omega_4$ form a basis of the corresponding ideal $\mathfrak{L}_1$ in $\mathfrak{G}_1$. If we set $\bar{\omega}_i = \sum_j \bar{a}_{ij}\bar{e}_j$, we have $N^2(\mathfrak{L}) = |a_{ij}| = |\bar{a}_{ij}|$.

Since a_{44} is prime to τ, ω_4 is not in $\mathfrak{G}_1$. Then by Lemmas 2, 3, $\mathfrak{L}_1$ is a principal ideal $\{\rho\}$, where ρ is an element in $\mathfrak{G}_1$. Then $\bar{e}_i\rho = \sum_j r_{ij}\bar{e}_j$ $(i = 1, \cdots, 4)$ form a basis of $\mathfrak{L}_1$, where the matrix (r_{ij}) is the transpose of the second matrix of ρ, with respect to the basis $\bar{e}_1, \cdots, \bar{e}_4$ of $\mathfrak{A}$.† Hence $|r_{ij}| = N^2(\rho)$. But the $\bar{\omega}_i$ also form a basis of $\mathfrak{L}_1$. Therefore $N^2(\mathfrak{L}) = |\bar{a}_{ij}| = N^2(\rho)$.

Since ρ is in $\mathfrak{L}$, $e_i\rho = \sum t_{ij}\omega_j = \sum_j s_{ij}e_j$, where the t's are rational integers and $(s_{ij}) = (t_{ij})(a_{ij})$ is the transpose of the second matrix of ρ with respect to the basis $e_1, \cdots, e_4$ of $\mathfrak{A}$. Hence $N^2(\rho) = |s_{ij}| = |t_{ij}| N^2(\mathfrak{L})$, $|t_{ij}| = 1$ and the $e_i\rho$ $(i = 1, \cdots, 4)$ form a basis of $\mathfrak{L}$. Therefore $\mathfrak{L}$ is the principal ideal defined by ρ.

It follows that every ideal in C is a principal ideal. But C was an arbitrarily chosen class. Hence every ideal in $\mathfrak{G}$ is principal. We have then by the second sentence in §1,

THEOREM 1. *Let $\mathfrak{A}$ be a rational generalized quaternion algebra with a negative fundamental number, and let $\mathfrak{G}$ be an arbitrarily chosen set of integral elements in $\mathfrak{A}$. Every one-sided ideal in $\mathfrak{G}$ is a principal ideal.*

It will be observed that in the above proof, the condition $N(\mathfrak{L}) \leqq 2(-d)^{1/2}$ was used only to insure a_{44} being prime to τ, i.e., that $\mathfrak{L}$ contains an element not in $\mathfrak{G}_1$. It follows from the above proof that if η is a non-singular element in the special $\mathfrak{G}$ considered above, there is a unit U in $\mathfrak{G}$, i.e., an element of norm ± 1, such that $U\eta = \rho$ is in $\mathfrak{G}_1$.‡

If we employ narrow equivalence of ideals, as in Tr., it may be shown that the number h_0, of classes of narrowly equivalent ideals in an arbitrary $\mathfrak{G}$ is independent of the particular $\mathfrak{G}$ which is considered; the proof of this being the same as Brandt's proof for the broader case.§ By the proof of Lemma 3, there is a single class of regular ideals in $\mathfrak{G}_1$, narrow equivalence being employed. Hence $\mathfrak{G}_1$ contains an element of norm -1. Therefore the same is true of the special $\mathfrak{G}$ employed in the proof of Theorem 1. Hence $h_0 = 1$ and every $\mathfrak{G}$ contains an element of norm -1.‖

---

* Brandt, loc. cit., p. 23.

† Dickson, *Algebras and their Arithmetics*, p. 95.

‡ Cf. Dickson, *Algebras and their Arithmetics*, p. 153, Lemma 7.

§ Brandt, loc. cit., p. 23.

‖ Cf. Brandt, loc. cit., p. 29.

5. **Applications of Theorem 1.** Let $\mathfrak{G}$ be as in Theorem 1 and let η_1, η_2 be elements of $\mathfrak{G}$ such that one of $N(\eta_1)$, $N(\eta_2)$, $\eta_1\eta_2'$ is $\neq 0$. Let $\mathfrak{L}$ be the set of all elements $\xi_1\eta_1+\xi_2\eta_2$, where ξ_1, ξ_2 are in $\mathfrak{G}$. The set $\mathfrak{L}$ is an ideal if it contains a non-singular element. This is obviously the case if either η_1 or η_2 is non-singular. Suppose they are both singular. For a properly chosen rational integer t, $\mathfrak{G}$ contains ti, tj, tij. We have $N(\xi\eta_1+\eta_2)=\xi\eta_1\eta_2'+\eta_2\eta_1'\xi'$, which is the trace of $\xi\eta_1\eta_2'$. Hence if we take $\xi=1$, ti, tj, or tij, $N(\xi\eta_1+\eta_2)\neq 0$ and $\mathfrak{L}$ is an ideal.

Let the g.c.r.d. and the g.c.l.d. of two elements in $\mathfrak{G}$ be defined as in Tr. Employing the above and Theorem 1, the following theorems may be proved by the same methods used in the proof of Theorem 4 of Tr.

THEOREM 2. *Let $\mathfrak{G}$ be as in Theorem* 1, *and let η_1, η_2 be elements of $\mathfrak{G}$ such that one of $N(\eta_1)$, $N(\eta_2)$, $\eta_1\eta_2'$ is $\neq 0$. Then η_1, η_2 have a* g.c.r.d. δ_1, *and a* g.c.l.d. δ_2 *and there are elements μ_1, μ_2, γ_1, γ_2 in $\mathfrak{G}$ such that $\mu_1\eta_1+\mu_2\eta_2=\delta_1$, $\eta_1\nu_1+\eta_2\nu_2=\delta_2$. Furthermore, δ_1 $[\delta_2]$ is uniquely determined, apart from a unit left [right] factor.*

THEOREM 3. *Let $\mathfrak{G}$ be as in Theorem* 1, *let λ be a non-singular element in $\mathfrak{G}$ not divisible by a rational prime, and let $N(\lambda)=\pm p_1\cdot p_2\cdot\cdots\cdot p_r$, where the p's are rational primes arranged in an arbitrary but fixed order. Then $\lambda=\pi_1\cdot\pi_2\cdot\cdots\cdot\pi_r$ where $N(\pi_i)=\pm p_i$ $(i=1, 2, \cdots, r)$ and each π_i is an element of $\mathfrak{G}$ which is uniquely determined apart from unit factors.*

We have seen that $\mathfrak{G}$ contains an element of norm -1. Hence, in the above factorization of λ, the sign of each $N(\pi_i)$, except one, may be arbitrarily chosen.

6. **An enlargement of a set of quadratic integers.** Let G be the set of all algebraic integers in the quadratic field defined by $(-\alpha)^{1/2}$, α being a rational integer with no square factor >1. Let β be a negative integer, with no square factor >1, such that $\alpha\equiv\beta \pmod 2$. Let $\mathfrak{A}$ be the rational generalized quaternion algebra with a basis $1, i, j, ij$ where $i^2=-\alpha$, $j^2=-\beta$, $ij=-ji$. Since $N(j)=\beta<0$, the fundamental number of $\mathfrak{A}$ is negative. By the results of Darkow or of the writer cited in the proof of Lemma 1, there is a set $\mathfrak{G}$ of integral elements in $\mathfrak{A}$ which contains a sub-set equivalent to G. Then the statements made in the last paragraph of §1 follow from Theorem 2.

UNIVERSITY OF KENTUCKY,
LEXINGTON, KY.

ON SIMPLY TRANSITIVE GROUPS WITH TRANSITIVE ABELIAN SUBGROUPS OF THE SAME DEGREE*

BY

DOROTHY MANNING

Introduction. It was in 1900 that William Burnside† first proved that a simply transitive permutation group of prime degree p must contain an invariant subgroup of order p. He later showed‡ that a simply transitive group of degree p^m with p a prime and $m>1$ is imprimitive and compound if it contains a permutation of order p^m. In 1921 he§ published a short proof that a simply transitive group in which there is a transitive Abelian subgroup of the same degree is compound unless all the operations of the Abelian subgroup are of the same prime order. However, the proof is not sound, as we shall see in §15.

In 1933 Professor I. Schur‖ published a paper on those simply transitive groups of degree n in which there is a transitive cyclic subgroup of degree n. He found them to be imprimitive and compound when n is not a prime. He seems to have been unaware of Burnside's paper of 1921. It was this fact that led W. A. Manning to a critical study of the status of the problem, and to the conviction that an investigation of the remaining cases should be undertaken.

In this paper it will be proved that *if a simply transitive group of degree* p^{a+b} (p *a prime and* $a\neq b$) *contains a transitive Abelian subgroup of the same degree and of type* (a, b), *it is imprimitive and compound.* That no such theorem holds when $a=b$ is shown by the existence of exceptions. There is an infinite family of simply transitive primitive groups¶ of degree k^2 ($k>2$) and order $2(k!)^2$ which contain transitive Abelian subgroups of degree k^2 that are

* Presented to the Society, April 6, 1935; received by the editors June 4, 1935.

† W. Burnside, *On the properties of groups of odd order*, Proceedings of the London Mathematical Society, vol. 33 (1900), p. 174; *On simply transitive groups of prime degree*, Quarterly Journal of Mathematics, vol. 37 (1906), p. 215; I. Schur, *Neuer Beweis eines Satzes von W. Burnside*, Jahresbericht der Deutschen Mathematiker-Vereinigung, vol. 17 (1908), p. 171; W. Burnside, *Theory of Groups of Finite Order*, 2d edition (1911), p. 341.

‡ W. Burnside, *Theory of Groups of Finite Order*, 2d edition (1911), p. 343.

§ W. Burnside, *On certain simply-transitive permutation-groups*, Proceedings of the Cambridge Philosophical Society, vol. 20 (1921), pp. 482–484.

‖ *Zur Theorie der einfach transitiven Permutationsgruppen*, Sitzungsberichte der Preussischen Akademie der Wissenschaften, Physikalisch-mathematische Klasse, 1933, pp. 598–623.

¶ W. A. Manning, *On the primitive groups of class ten*, American Journal of Mathematics, vol. 28 (1906), p. 235.

direct products of two Abelian subgroups of order k. The members of this family with $k=p^a>2$ are exceptions to the proposed theorem.

There is a somewhat similar problem for solvable groups, first encountered by Professor J. F. Ritt.* He proved that a primitive solvable group in p^m letters with p prime and $m>1$ cannot contain a substitution of order p^m except in the case $p=2$, $m=2$.

It should be noted that if an imprimitive group contains a transitive Abelian subgroup of the same degree, it is compound.† For if it were simple, the group according to which its systems of imprimitivity are permuted would contain a transitive Abelian subgroup of order n and degree less than n, an impossibility, since a transitive Abelian group is always regular.

1. In this section we reproduce without proof as much of the preliminary work of Burnside's paper of 1921‡ as we shall need.

Let a and b be natural numbers, p and q primes, the same or different. Any transitive Abelian permutation group of degree $n=p^aq^b$ generated by two permutations of orders p^a and q^b, may be written $\{P, Q\}$, where

$$\left.\begin{array}{ll} P: & x_{u,v} = x'_{u+1,v} \\ Q: & x_{u,v} = x'_{u,v+1} \end{array}\right\} (u = 0, 1, \cdots, p^a - 1; v = 0, 1, \cdots, q^b - 1),$$

the two symbols $x_{u,v}$ and $x_{y,z}$ representing the same letter if and only if $u\equiv y$ mod p^a and $v\equiv z$ mod q^b. (This rule is also to be applied to the $\xi_{u,v}$, the $\Gamma_{u,v}$, the $\Delta_{u,v}$, the $M_{u,v}$, the $m_{u,v}$, the $T_{i,j}$, and the $W_{i,j}$, which will be introduced later.)

Let $H=\{P, Q\}$ and let G be any simply transitive permutation group containing H as a proper subgroup.

If ϵ is a primitive p^ath and η a primitive q^bth root of unity,

$$T: \quad x_{u,v} = \frac{1}{n} \sum_{y=0}^{p^a-1} \sum_{z=0}^{q^b-1} \epsilon^{uy}\eta^{vz}\xi_{y,z} \quad (u = 0, 1, \cdots, p^a - 1; v = 0, 1, \cdots, q^b - 1)$$

completely reduces both G and H.

$$\left.\begin{array}{ll} (1.1) & T^{-1}PT: \ \xi_{y,z} = \epsilon^{y}\xi'_{y,z} \\ (1.2) & T^{-1}QT: \ \xi_{y,z} = \eta^{z}\xi'_{y,z} \end{array}\right\} (y = 0, 1, \cdots, p^a - 1; z = 0, 1, \cdots, q^b - 1).$$

---

* *On algebraic functions which can be expressed in terms of radicals*, these Transactions, vol. 24 (1922), p. 27.

† This has already been proved by Professor Schur when the Abelian subgroup is cyclic, *Zur Theorie der einfach transitiven Permutationsgruppen*, Sitzungsberichte der Preussischen Akademie der Wissenschaften, Physikalisch-mathematische Klasse, 1933, p. 622.

‡ Proceedings of the Cambridge Philosophical Society, vol. 20 (1921), pp. 482–483. The mistakes occur on page 484.

If we represent the variables of an irreducible constituent group Γ of $T^{-1}GT$: ξ_{α_v,β_v} $(v=1, 2, \cdots, \mu)$ (where μ is the degree of Γ), and the letters of a transitive constituent Δ of G_0 (the subgroup of G which fixes $x_{0,0}$): x_{a_u,b_u} $(u=1, 2, \cdots, m)$ (where m is the degree of Δ), and if $\chi_{P^{a_u}Q^{b_u}}$ represents the trace of the substitution of Γ which corresponds to $P^{a_u}Q^{b_u}$, then

$$\chi_{P^{a_u}Q^{b_u}} = \sum_{v=1}^{\mu} \epsilon^{a_u\alpha_v}\eta^{b_u\beta_v} = \frac{\mu}{m}\sum_{u=1}^{m} \epsilon^{a_u\alpha_v}\eta^{b_u\beta_v}, \tag{1.3}$$

provided Γ is not the identical constituent (on $\xi_{0,0}$ alone) and Δ is not the transitive constituent of G_0 on $x_{0,0}$.

2. We now establish a sufficient condition that the group G be imprimitive.

LEMMA I. *If any permutation S of G which is not also in G_0 corresponds to the identity of any irreducible constituent Γ of $T^{-1}GT$ except the identical constituent, G is imprimitive.**

The correspondence between G and Γ is not one to one, for both S and the identity of G correspond to the identity of Γ. Let Γ' be the subgroup of Γ which corresponds to G_0. Let G' be the largest (proper or improper) subgroup of G which corresponds to Γ'. Then G_0 is a proper subgroup of G', for G' contains S. Since G is simply transitive, $\xi_{\alpha_1,\beta_1}+\cdots+\xi_{\alpha_\mu,\beta_\mu}$ is an absolute invariant of Γ',† while Γ, being irreducible, can have none. Hence Γ' is a proper subgroup of Γ, and G' is a proper subgroup of G. Therefore, G_0 is a non-maximal subgroup of G, and G is imprimitive.

3. Throughout this paper, i and j will represent positive integers relatively prime to p and q, respectively.

We introduce

$$T_{i,j}: \qquad x_{u,v} = \frac{1}{n}\sum_{y=0}^{p^a-1}\sum_{z=0}^{q^b-1}(\epsilon^i)^{uy}(\eta^j)^{vz}\xi_{y,z}$$
$$(u = 0, 1, \cdots, p^a-1;\ v = 0, 1, \cdots, q^b-1),$$

and let $\Gamma_{y,z}$ represent the irreducible constituent of $T^{-1}GT$ which affects, among others, the variable $\xi_{y,z}$, and let $\Gamma_{y,z}^{(i,j)}$ represent the irreducible constituent of $T_{i,j}^{-1}GT_{i,j}$ which affects $\xi_{y,z}$. If we replace ϵ by ϵ^i and η by η^j, we obtain from the substitution $T^{-1}AT$, the substitution $T_{i,j}^{-1}AT_{i,j}$.

Since we can write T in the form

---

* Cf. W. Burnside, *Theory of Groups of Finite Order*, 2d edition (1911), p. 342.

† W. Burnside, *On certain simply-transitive permutation-groups*, Proceedings of the Cambridge Philosophical Society, vol. 20 (1921), p. 483.

$$x_{u,v} = \frac{1}{n} \sum_{y=0}^{p^a-1} \sum_{z=0}^{q^b-1} \epsilon^{u(iy)} \eta^{v(jz)} \xi_{iy,jz}$$
$$(u = 0, 1, \cdots, p^a - 1; v = 0, 1, \cdots, q^b - 1),$$

$T = T_{i,j} W_{i,j}$, where $W_{i,j}$ is the permutation $\xi_{y,z} = \xi'_{iy,jz}$ ($y = 0, 1, \cdots, p^a - 1$; $z = 0, 1, \cdots, q^b - 1$). Hence $T^{-1}AT = W_{i,j}^{-1}(T_{i,j}^{-1} A T_{i,j}) W_{i,j}$. This shows that $\Gamma_{y,z}^{(i,j)}$ and $\Gamma_{iy,jz}$ differ only in the names of their variables. That is, if we consider the corresponding matrix groups, the variables of $\Gamma_{y,z}^{(i,j)}$ being denoted $\xi_{\alpha_1,\beta_1}, \xi_{\alpha_2,\beta_2}, \cdots, \xi_{\alpha_\mu,\beta_\mu}$ and the rows and columns of the matrices of the first matrix group being made to correspond to the variables in that order, then the variables of $\Gamma_{iy,jz}$ are $\xi_{i\alpha_1,j\beta_1}, \xi_{i\alpha_2,j\beta_2}, \cdots, \xi_{i\alpha_\mu,j\beta_\mu}$ and if the rows and columns of the matrices of the second group are made to correspond to the variables in this order, then the two matrix groups consist of the same matrices and the two matrices which correspond to A are equal.* Combining the results of these two paragraphs, we have

LEMMA II. *If we replace ϵ by ϵ^i and η by η^j, we "replace $\Gamma_{y,z}$ by $\Gamma_{iy,jz}$," that is, from the matrix of $\Gamma_{y,z}$ which corresponds to any permutation of G, we obtain the matrix of $\Gamma_{iy,jz}$ which corresponds to the same permutation of G, provided that if the rows and columns of the matrices of $\Gamma_{y,z}$ correspond to the variables $\xi_{\alpha_1,\beta_1}, \xi_{\alpha_2,\beta_2}, \cdots, \xi_{\alpha_\mu,\beta_\mu}$ in that order, then the rows and columns of the matrices of $\Gamma_{iy,jz}$ (which necessarily correspond to the variables $\xi_{i\alpha_1,j\beta_1}, \xi_{i\alpha_2,j\beta_2}, \cdots, \xi_{i\alpha_\mu,j\beta_\mu}$) correspond to them in that order.*

This gives us

LEMMA III. *If we replace ϵ by ϵ^i and η by itself, we replace $\Gamma_{0,z}$ by itself. If we replace ϵ by itself and η by η^j, we replace $\Gamma_{y,0}$ by itself.*

4. The operation of replacing ϵ by ϵ^i and η by η^j, we may denote by

$$\begin{pmatrix} \epsilon & , & \eta \\ \epsilon^i & , & \eta^j \end{pmatrix}.$$

These operations form a group K which is simply isomorphic to the group K' whose letters are the $p^a q^b$ symbols $\Gamma_{y,z}$ and whose permutations are of the form

$$\begin{pmatrix} \Gamma_{y,z} \\ \Gamma_{iy,jz} \end{pmatrix} = \begin{pmatrix} \Gamma_{0,0} & \cdots & \Gamma_{p^a-1,q^b-1} \\ \Gamma_{0,0} & \cdots & \Gamma_{i(p^a-1),j(q^b-1)} \end{pmatrix}.$$

Its order is $\phi(p^a)\phi(q^b)$.†

* Two matrices are equal to one another if and only if corresponding elements are equal.

† Euler's ϕ-function.

Our system of notation for the irreducible constituents gives us as many different symbols for a given irreducible constituent as it has variables. The actual number γ of irreducible constituents is less than $p^a q^b$, for otherwise G would be Abelian and H would not be a proper subgroup of G. Let K'' represent the group according to which the irreducible constituents are permuted by the operations of K. Then K'' is of degree γ, its order lies between 1 and $\phi(p^a)\phi(q^b)$, inclusive, and the number of its transitive constituents (counting those on one letter) lies between 2 and $(a+1)(b+1)$, inclusive, and will be shown later to be greater than 2.

DEFINITION. *The $\phi(p^{a-u})\phi(q^{b-v})$ variables ξ_{ip^u,jq^v} for given values of u and v $(0 \leqq u \leqq a,\ 0 \leqq v \leqq b)$ will be said to belong to the (u, v) category.*

LEMMA IV. *The variables of the (u, v) category are affected by irreducible constituents all of which belong to the same transitive constituent of K''.*

For Γ_{ip^u,jq^v} is replaced by $\Gamma_{i'p^u,j'q^v}$ by the operation

$$\begin{pmatrix} \epsilon & , & \eta \\ \epsilon^i & , & \eta^j \end{pmatrix}^{-1} \cdot \begin{pmatrix} \epsilon & , & \eta \\ \epsilon^{i'} & , & \eta^{j'} \end{pmatrix}$$

of K.

We shall say that the (u, v) category belongs to this transitive constituent of K''.

It follows from Lemma II that all the irreducible constituents of a given transitive constituent of K'' have the same number of variables belonging to any given category. Hence Lemma IV gives us

LEMMA V. *If there are t irreducible constituents of $T^{-1}GT$ in a given transitive constituent of K'' and if one of these irreducible constituents affects a variable of the (u, v) category, then it must affect exactly $1/t$ of the variables of that category.*

5. We now prove

LEMMA VI. *If $\Gamma_{y,0}$ affects ξ_{ip^u,jq^v} $(v<b)$, it affects each of the variables ξ_{ip^u,q^v}, $\xi_{ip^u,2q^v}$, $\cdots$, $\xi_{ip^u,(q-1)q^v}$, $\xi_{ip^u,(q+1)q^v}$, $\cdots$, ξ_{ip^u,q^b-q^v}. Whether $v<b$ or not, the number t of irreducible constituents in the transitive constituent of K'' to which $\Gamma_{y,0}$ belongs must divide $\phi(p^{a-u})$. Similarly for $\Gamma_{0,z}$. If $\Gamma_{y,0}=\Gamma_{0,z}$, $t=1$ and this irreducible constituent affects all the variables of the (u, v) category.*

If $\Gamma_{y,0}$ affects ξ_{ip^u,jq^v}, $\Gamma_{y,0}=\Gamma_{ip^u,jq^v}$. By Lemma III, if we replace ϵ by itself and η by any primitive q^bth root of unity, $\eta^{j'}$ (say), we replace $\Gamma_{y,0}$ by itself. By Lemma II, if we do the same thing to Γ_{ip^u,jq^v}, we obtain $\Gamma_{ip^u,jj'q^v}$. If we let j' run through a reduced set of residues mod q^b, jj' will do the same, while if $v<b$, $jj'q^v$ will assume values congruent to q^v, $2q^v$, $\cdots$, $(q-1)q^v$,

$(q+1)q^v, \cdots, q^b - q^v \bmod q^b$, each $\phi(q^b)/\phi(q^{b-v})$ times. Hence the first statement.

The second statement then follows from Lemma V.

The irreducible constituent $\Gamma_{v,0}$ is invariant under the operation

$$\begin{pmatrix} \epsilon & , & \eta \\ \epsilon & , & \eta^i \end{pmatrix},$$

and $\Gamma_{0,z}$ is invariant under

$$\begin{pmatrix} \epsilon & , & \eta \\ \epsilon^i & , & \eta \end{pmatrix}.$$

Hence if $\Gamma_{v,0} = \Gamma_{0,z}$, it is invariant under their product

$$\begin{pmatrix} \epsilon & , & \eta \\ \epsilon^i & , & \eta^i \end{pmatrix},$$

and thus under the group K. Then $\Gamma_{v,0}$ alone forms one of the transitive constituents of K'' and, by Lemma V, it must affect all the variables of each category which belongs to it.

LEMMA VII. *The group K'' has at least two transitive constituents besides the one on $\Gamma_{0,0}$ alone.*

If K'' had only one transitive constituent besides the one on $\Gamma_{0,0}$ alone, all the variables except $\xi_{0,0}$ would belong to it. Then each of its irreducible constituents would contain at least one variable of an (a, v) category with $v < b$ and at least one variable of a (u, b) category with $u < a$. Hence, by Lemma VI, it would consist of just one irreducible constituent. Then $T^{-1}GT$ would consist of two irreducible constituents. But that is impossible, for G is, by hypothesis (§1), simply transitive.*

6. We now restrict our attention to the irreducible constituent $\Gamma_{p^{a-1},0}$ and the permutations P^vQ^z of the Abelian subgroup H. Let $\chi_{P^vQ^z}$ represent the trace of the substitution of $\Gamma_{p^{a-1},0}$ which corresponds to P^vQ^z and let μ represent the degree of $\Gamma_{p^{a-1},0}$.

It follows from Lemma VI that for the transitive constituent of K'' to which $\Gamma_{p^{a-1},0}$ belongs, t must divide $p-1$. Let

$$c = (p-1)/t. \tag{6.1}$$

It also follows from Lemma VI (together with equations (1.1) and (1.2)) that we may write

* W. Burnside, *Theory of Groups of Finite Order*, 2d edition (1911), p. 339.

$$
\begin{aligned}
\chi_{PQ} = {} & d_{a,b}(\epsilon^{A_1} + \cdots + \epsilon^{A_{cp^{a-1}}})(\eta + \cdots + \eta^{q-1} + \eta^{q+1} + \cdots + \eta^{q^b-1}) \\
& + d_{a-1,b}(\epsilon^{pB_1} + \cdots + \epsilon^{pB_{cp^{a-2}}})(\eta + \cdots + \eta^{q-1} + \eta^{q+1} + \cdots + \eta^{q^b-1}) + \cdots \\
& + d_{1,b}(\epsilon^{p^{a-1}C_1} + \cdots + \epsilon^{p^{a-1}C_c})(\eta + \cdots + \eta^{q-1} + \eta^{q+1} + \cdots + \eta^{q^b-1}) \\
& \qquad + d_{0,b}(\eta + \cdots + \eta^{q-1} + \eta^{q+1} + \cdots + \eta^{q^b-1}) \\
& + d_{a,b-1}(\epsilon^{D_1} + \cdots + \epsilon^{D_{cp^{a-1}}})(\eta^q + \cdots + \eta^{(q-1)q} + \eta^{(q+1)q} + \cdots + \eta^{q^b-q}) \\
& \qquad + \cdots + d_{0,b-1}(\eta^q + \cdots + \eta^{(q-1)q} + \eta^{(q+1)q} + \cdots + \eta^{q^b-q}) + \\
& \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \\
& + d_{a,1}(\epsilon^{E_1} + \cdots + \epsilon^{E_{cp^{a-1}}})(\eta^{q^{b-1}} + \cdots + \eta^{q^b-q^{b-1}}) + \cdots \\
& \qquad + d_{0,1}(\eta^{q^{b-1}} + \cdots + \eta^{q^b-q^{b-1}}) \\
& + d_{a,0}(\epsilon^{F_1} + \cdots + \epsilon^{F_{cp^{a-1}}}) + d_{a-1,0}(\epsilon^{pG_1} + \cdots + \epsilon^{pG_{cp^{a-2}}}) + \cdots \\
& \qquad + d_{1,0}(\epsilon^{p^{a-1}H_1} + \cdots + \epsilon^{p^{a-1}H_c}) + d_{0,0},
\end{aligned}
\tag{6.2}
$$

where $d_{k,l}=1$ or 0 depending on whether the variables of the $(a-k,\ b-l)$ category do or do not belong to the transitive constituent of K'' which includes $\Gamma_{p^{a-1},0}$, and where each $A, B, \cdots, H$ is relatively prime to p and no two A's are equal, no two B's are equal, etc. Notice that $d_{1,0}=1$ and $d_{0,0}=0$.

Because of (1.1) and (1.2), we have

$$
T^{-1}P^vQ^zT: \qquad \xi_{u,v} = \epsilon^{vu}\eta^{zv}\xi'_{u,v} \qquad (u = 0, 1, \cdots, p^a - 1;\ v = 0, 1, \cdots, q^b - 1). \tag{6.3}
$$

This shows that the substitution of $\Gamma_{p^{a-1},0}$ which corresponds to P^vQ^z is a multiplication and that it may be obtained from the substitution of $\Gamma_{p^{a-1},0}$ which corresponds to PQ by replacing ϵ by ϵ^v and η by η^z in each multiplier. We may obtain $\chi_{P^vQ^z}$ from χ_{PQ} by replacing ϵ by ϵ^v and η by η^z provided χ_{PQ} is written out in full as the sum of the μ multipliers just as they are taken from (6.3). (Any cancellation or substitution making use of any relation existing between powers of ϵ or powers of η, or between ϵ and η, etc., before replacing ϵ by ϵ^v and η by η^z will in general lead to incorrect results.) However, the expression in the right-hand member of (6.2) is such that we may obtain $\chi_{P^vQ^z}$ from it by replacing ϵ by ϵ^v and η by η^z.

We now make use of the following procedure: Replace ϵ by ϵ^v and η by η^z in the right-hand member of (6.2) (obtaining $\chi_{P^vQ^z}$). The sums obtained from the sums of powers of η are equal to ordinary integers. Replace each by the corresponding integer and collect like powers of ϵ, reducing exponents mod p^a. The result is a polynomial in ϵ with integers as coefficients. We write it

$$
\chi_{P^vQ^z} = \sigma_0 + \sigma_1\epsilon + \sigma_2\epsilon^2 + \cdots + \sigma_{p^a-1}\epsilon^{p^a-1}.
$$

Using the first and third members of (1.3), we obtain

$$\sum_{k=0}^{p^a-1} \sigma_k \epsilon^k = \frac{\mu}{m} \sum_{u=1}^{m} \epsilon^{a_u \cdot p^{a-1}} \eta^{b_u \cdot 0}, \tag{6.4}$$

where m denotes the degree of $\Delta_{y,z}$, the transitive constituent of G_0 which includes $x_{y,z}$ and where $x_{a_1,b_1}, x_{a_2,b_2}, \cdots, x_{a_m,b_m}$ are the letters of $\Delta_{y,z}$. Equation (6.4) holds except when $y \equiv 0 \bmod p^a$ and $z \equiv 0 \bmod q^b$ at the same time. From (6.4) we obtain

$$\sum_{k=0}^{p^a-1} \sigma_k \epsilon^k = \frac{\mu}{m} \sum_{l=0}^{p-1} \pi_l \epsilon^{l p^{a-1}},$$

where π_l is the number of letters x_{a_u,b_u} of $\Delta_{y,z}$ whose first subscripts a_u are congruent to $l \bmod p$. Hence ϵ is a root of the equation

$$\sum_{k=0}^{p^a-1} \sigma_k x^k - \frac{\mu}{m} \sum_{l=0}^{p-1} \pi_l x^{l p^{a-1}} = 0, \tag{6.5}$$

whose coefficients are rational numbers. But ϵ is also a root of the equation

$$1 + x^{p^{a-1}} + x^{2p^{a-1}} + \cdots + x^{(p-1)p^{a-1}} = 0, \tag{6.6}$$

which is irreducible in the field of rational numbers.* Hence if the left-hand member of (6.5) is not identically zero, it must be divisible by the left-hand member of (6.6). In either case we may write the left-hand member of (6.5) in the form

$$(\gamma_0 + \gamma_1 x + \cdots + \gamma_{p^{a-1}-1} x^{p^{a-1}-1})(1 + x^{p^{a-1}} + x^{2p^{a-1}} + \cdots + x^{(p-1)p^{a-1}}).$$

Therefore,

$$\sigma_0 - \frac{\mu}{m}\pi_0 = \sigma_{p^{a-1}} - \frac{\mu}{m}\pi_1 = \sigma_{2p^{a-1}} - \frac{\mu}{m}\pi_2 = \cdots = \sigma_{(p-1)p^{a-1}} - \frac{\mu}{m}\pi_{p-1},$$

and

$$m = \sum_{l=0}^{p-1} \pi_l = \sum_{l=0}^{p-1} \left\{ \pi_k + \frac{m}{\mu} (\sigma_{l p^{a-1}} - \sigma_{k p^{a-1}}) \right\}.$$

Hence

$$m\{\mu + p\sigma_{k p^{a-1}} - (\sigma_0 + \sigma_{p^{a-1}} + \cdots + \sigma_{(p-1)p^{a-1}})\} = p\mu\pi_k \quad (k = 0, 1, \cdots, p-1). \tag{6.7}$$

Let

$$M_{y,z} = (p-1)\sigma_0 - (\sigma_{p^{a-1}} + \sigma_{2p^{a-1}} + \cdots + \sigma_{(p-1)p^{a-1}}), \tag{6.8}$$

* Weber, *Lehrbuch der Algebra*, 2d edition (1898), vol. 1, pp. 459 and 600.

for the permutation P^yQ^z, and write $m_{y,z}$ instead of m for the degree of $\Delta_{y,z}$. When $k=0$, (6.7) becomes

$$m_{y,z}\{\mu + M_{y,z}\} = p\mu\pi_0. \tag{6.9}$$

Adding the $p-1$ equations for which $k=1, 2, \cdots, p-1$, we obtain

$$m_{y,z}\{(p-1)\mu - M_{y,z}\} = p\mu(\pi_1 + \pi_2 + \cdots + \pi_{p-1}). \tag{6.10}$$

When using these formulas, we must remember that $\pi_0, \cdots, \pi_{p-1}$ depend on $\Delta_{y,z}$ and that these formulas do not apply to $\Delta_{0,0}$.

7. Since p and μ are constants and since $m_{y,z}$ and π_0 can each have but one value for a given transitive constituent of G_0, (6.9) gives us

LEMMA VIII. *If $\Delta_{y',z'}$ is the same transitive constituent of G_0 as $\Delta_{y,z}$, then $M_{y',z'}=M_{y,z}$.*

DEFINITION. *The $\phi(p^{a-u})\phi(q^{b-v})$ letters x_{ip^u,jq^v} (u and v fixed) will be said to belong to the (u, v) category of letters.*

From the meaning of the π_l it follows that $\Delta_{y,z}$ includes at least one letter of some $(0, v)$ category if and only if $\pi_1+\pi_2+\cdots+\pi_{p-1}>0$. Since $m_{y,z}$, μ, and p are positive integers, (6.10) shows us that when $\Delta_{y,z}$ is not $\Delta_{0,0}$, $\pi_1+\pi_2+\cdots+\pi_{p-1}>0$ if and only if $(p-1)\mu-M_{y,z}>0$.

LEMMA IX. *If $(p-1)\mu-M_{y,z}>0$ and $\Delta_{y,z}$ is not $\Delta_{0,0}$, $\Delta_{y,z}$ includes at least one $x_{i,v}$ and $M_{y,z}$ is equal to at least one of the $M_{i,v}$.*

8. We now develop two sufficient conditions for the imprimitivity of G.

LEMMA X. *If an irreducible constituent of $T^{-1}GT$ which is not $\Gamma_{0,0}$ affects no variable of any $(0, v)$ category (or if it affects no variable of any $(u, 0)$ category), G is imprimitive.*

For if Γ affects no variable of any $(0, v)$ category, (1.1) shows that $P^{p^{a-1}}$ corresponds to the identity of Γ, and then, by Lemma I, G is imprimitive. If Γ affects no variable of any $(u, 0)$ category, $Q^{q^{b-1}}$ corresponds to the identity of Γ and again G is imprimitive.

LEMMA XI. *If $d_{a,b}=d_{a-1,b}=\cdots=d_{0,b}$, or if $d_{a,b}=d_{a,b-1}=\cdots=d_{a,0}$, G is imprimitive.*

For if $d_{a,b}=d_{a-1,b}=\cdots=d_{0,b}=0$, it follows from the meaning of the $d_{k,l}$ (§6) that $\Gamma_{p^{a-1},0}$ contains no variable of the (0, 0), the (1, 0), $\cdots$, or the $(a, 0)$ category, and then, by Lemma X, G is imprimitive. If $d_{a,b}=d_{a-1,b}=\cdots=d_{0,b}=1$, all the variables of those categories belong to the transitive constituent of K'' which includes $\Gamma_{p^{a-1},0}$. By Lemma VII, there must then be at least one transitive constituent of K'' besides the one

on $\Gamma_{0,0}$ alone whose irreducible constituents affect none of the variables of those categories. We apply Lemma X to one of those irreducible constituents and see that G is imprimitive. Hence, if $d_{a,b}=d_{a-1,b}=\cdots=d_{0,b}$, G is imprimitive. The proof of the other half of the lemma is similar.

9. To simplify the formulas, we let

$$d_{a+1,0}=\cdots=d_{a+1,b}=d_{0,b+1}=\cdots=d_{a,b+1}=d_{a+1,b+1}=0, \tag{9.1}$$

and let $R_{u,v}$ represent the number of variables in the $(a-u,\ b-v)$ category affected by $\Gamma_{p^{a-1},0}$. Then

$$R_{u,v}=d_{u,v}\phi(p^u)\phi(q^v)/t. \tag{9.2}$$

Equation (6.8) now gives us

$$\begin{aligned} M_{ip^k,jq^l} &= d_{k+1,l+1}cp^kq^l-\sum_{v=0}^{l}R_{k+1,v} \\ &\quad+(p-1)\Big(-\sum_{u=1}^{k}d_{u,l+1}cp^{u-1}q^l-d_{0,l+1}q^l+\sum_{u=0}^{k}\sum_{v=0}^{l}R_{u,v}\Big). \end{aligned} \tag{9.3}$$

(To obtain this expression, it is neither necessary nor desirable to calculate $\sigma_{p^{a-1}}$, $\sigma_{2p^{a-1}}$, $\cdots$, $\sigma_{(p-1)p^{a-1}}$ separately. It is easier to calculate their sum directly.)

To calculate $(p-1)\mu-M_{ip^k,jq^l}$, we make use of the fact that we may write

$$\begin{aligned} \mu=\sum_{u=0}^{a}\sum_{v=0}^{b}R_{u,v} &= \sum_{u=0}^{k}\sum_{v=0}^{l}R_{u,v}+\sum_{u=0}^{k}\sum_{v=l+1}^{l+1}R_{u,v}+\sum_{u=k+1}^{k+1}\sum_{v=0}^{l}R_{u,v} \\ &\quad+\sum_{u=k+1}^{k+1}\sum_{v=l+1}^{l+1}R_{u,v}+\sum_{u=k+2}^{a}\sum_{v=0}^{l+1}R_{u,v}+\sum_{u=0}^{k+1}\sum_{v=l+2}^{b}R_{u,v}+\sum_{u=k+2}^{a}\sum_{v=l+2}^{b}R_{u,v}, \end{aligned}$$

where any summation is to be considered equal to zero if one or both of its $\sum$'s starts with a number larger than the one with which it ends. Then

$$\begin{aligned} (p-1)\mu-M_{ip^k,jq^l} &= (p-1)\Big\{\sum_{u=k+2}^{a}\sum_{v=0}^{l+1}R_{u,v}+\sum_{u=0}^{k+1}\sum_{v=l+2}^{b}R_{u,v} \\ &\quad+\sum_{u=k+2}^{a}\sum_{v=l+2}^{b}R_{u,v}\Big\}+d_{k+1,l+1}cp^kq^l\{(p-1)(q-1)-1\} \\ &\quad+\sum_{v=1}^{l}d_{k+1,v}cp^{k+1}(q-1)q^{v-1}+d_{k+1,0}cp^{k+1} \\ &\quad+\sum_{u=1}^{k}d_{u,l+1}c(p-1)p^{u-1}q^{l+1}+d_{0,l+1}(p-1)q^{l+1}. \end{aligned} \tag{9.4}$$

10. Since we wish to prove whenever it may be possible that G is imprimitive, we need only consider the case when the conditions of Lemma XI are not satisfied. Then at least one of the $d_{u,b}=1$ and at least one of the $d_{a,v}=1$. We seek to show, when possible, that under these conditions $(p-1)\mu-M_{ip^k,jq^l}>0$.

Notice first that each term of (9.4) is $\geqq 0$, and that each term is >0 if the corresponding $d_{u,v}\neq 0$ with the exception of the term $d_{k+1,l+1}cp^kq^l\{(p-1)(q-1)-1\}$ when $p=q=2$.

If $k+2\leqq a$, the terms which involve the $d_{a,v}$ $(v=0, 1, \cdots, b)$ occur within the first pair of brackets. Hence, since at least one of the $d_{a,v}=1$, $(p-1)\mu-M_{ip^k,jq^l}>0$. Similarly if $l+2\leqq b$.

If $k=a-1$ and $l=b-1$, the $d_{a,v}$ $(v=0, 1, \cdots, b)$ and $d_{u,b}$ $(u=0, 1, \cdots, a-1)$ occur in the last five terms (counting a summation as a term now) and $(p-1)\mu-M_{ip^k,jq^l}>0$ except when $d_{a,0}=\cdots=d_{a,b-1}=d_{0,b}=\cdots=d_{a-1,b}=0$ and $p=q=2$, and then it equals zero.

If $k=a-1$ and $l=b$, the $d_{a,v}$ $(v=0, 1, \cdots, b)$ occur in the terms $\sum_{v=1}^{l}d_{k+1,v}cp^{k+1}(q-1)q^{v-1}$ and $d_{k+1,0}cp^{k+1}$. Hence $(p-1)\mu-M_{ip^k,jq^l}>0$. Similarly if $k=a$ and $l=b-1$.

The case $k=a$, $l=b$ corresponds to $\Delta_{0,0}$ and is of no interest since we cannot apply Lemma IX.

Combining these results, we find

LEMMA XII. *Unless one or more of the three conditions*

(1) $d_{a,b}=d_{a,b-1}=\cdots=d_{a,0}$,

(2) $d_{a,b}=d_{a-1,b}=\cdots=d_{0,b}$, *or*

(3) $p=q=2$

holds, $(p-1)\mu-M_{ip^k,jq^l}>0$ *for each transitive constituent* Δ_{ip^k,jq^l} *not* $\Delta_{0,0}$.

When conditions (1) and (2) do not hold but condition (3) does, $(p-1)\mu-M_{ip^k,jq^l}=\mu-M_{2^k,2^l}$. Now $\mu-M_{2^k,2^l}>0$ for all transitive constituents (not $\Delta_{0,0}$) except perhaps $\Delta_{2^{a-1},2^{b-1}}$. Lemma IX and its proof then tell us that when $\mu-M_{2^{a-1},2^{b-1}}=0$, each transitive constituent of G_0 includes at least one letter of some $(0, v)$ category except $\Delta_{0,0}$ and $\Delta_{2^{a-1},2^{b-1}}$, and that they do not. Then $x_{2^{a-1},2^{b-1}}$ must be the only letter of $\Delta_{2^{a-1},2^{b-1}}$, and G_0 fixes at least two letters, $x_{0,0}$ and $x_{2^{a-1},2^{b-1}}$. Hence G is imprimitive.

Therefore, regardless of whether $p=q=2$ or not, we have, with the aid of Lemmas XI and IX,

LEMMA XIII. *If G is primitive, each transitive constituent of G_0 except $\Delta_{0,0}$ includes at least one letter of some $(0, v)$ category, and M_{p^k,q^l} is equal to at least one of the M_{i,jq^h} when $k+l<a+b$. Also each transitive constituent includes at least one letter of some $(u, 0)$ category, and M_{p^k,q^l} is equal to at least one of the $M_{ip^l,j}$.*

11. We can obtain an expression for M_{ip^k,jq^l} which is more convenient than (9.3) for comparing the M's as indicated in Lemma XIII. To this end, we observe that $(p-1)d_{0,v}=cd_{0,v}$, for if $d_{0,v}=0$, it is obviously true and if $d_{0,v}=1$, $\Gamma_{p^{a-1},0}=\Gamma_{0,jq^{b-v}}$, $t=1$ (Lemma VI), and (6.1) gives us $c=p-1$. Starting from (9.3), we replace each $R_{u,v}$ by its value in terms of $d_{u,v}$, c, p, q, u, and v, making use of (9.2) and (6.1). The coefficient of each $d_{0,v}$ contains $p-1$ as a factor and we replace $(p-1)d_{0,v}$ by $cd_{0,v}$. On simplifying, we find

$$M_{ip^k,jq^l} = c\sum_{u=0}^{k}\sum_{v=0}^{l} D_{u,v}p^uq^v, \tag{11.1}$$

where

$$D_{u,v} = d_{u,v} - d_{u,v+1} - d_{u+1,v} + d_{u+1,v+1}. \tag{11.2}$$

LEMMA XIV. *$D_{u,v}$ is an integer and $|D_{u,v}|\leqq 2$. Furthermore, $|D_{a,v}|\leqq 1$ and $|D_{u,b}|\leqq 1$.*

For each $d_{u,v}$ is equal to one or zero (§6) and each $d_{a+1,v}$ and $d_{u,b+1}$ is zero (§9).

LEMMA XV. *If $M_{p^k,q^l}=M_{i,jq^h}$ with $h<l$,*

$$\sum_{u=1}^{k}\sum_{v=0}^{l} D_{u,v}p^uq^v = -\sum_{v=h+1}^{l} D_{0,v}q^v;$$

with $h=l$,

$$\sum_{u=1}^{k}\sum_{v=0}^{l} D_{u,v}p^uq^v = 0;$$

with $h>l$,

$$\sum_{u=1}^{k}\sum_{v=0}^{l} D_{u,v}p^uq^v = \sum_{v=l+1}^{h} D_{0,v}q^v.$$

If $M_{p^k,q^l}=M_{ip^f,j}$ with $f<k$,

$$\sum_{u=0}^{k}\sum_{v=1}^{l} D_{u,v}p^uq^v = -\sum_{u=f+1}^{k} D_{u,0}p^u;$$

with $f=k$,

$$\sum_{u=0}^{k}\sum_{v=1}^{l} D_{u,v}p^uq^v = 0;$$

with $f>k$,

$$\sum_{u=0}^{k}\sum_{v=1}^{l} D_{u,v}p^uq^v = \sum_{u=k+1}^{f} D_{u,0}p^u.$$

These may be obtained from (11.1).

LEMMA XVI. *If* $D_{a,0}=D_{a,1}=\cdots=D_{a,b-1}=0$, *or if* $D_{0,b}=D_{1,b}=\cdots=D_{a-1,b}=0$, G *is imprimitive.*

For if $D_{a,0}=D_{a,1}=\cdots=D_{a,b-1}=0$, $d_{a,0}=d_{a,1}=\cdots=d_{a,b-1}=d_{a,b}$ and if $D_{0,b}=D_{1,b}=\cdots=D_{a-1,b}=0$, $d_{0,b}=d_{1,b}=\cdots=d_{a-1,b}=d_{a,b}$. In either case G is imprimitive, by Lemma XI.

For future convenience we introduce the function $F(h, l)$ defined in this way:

$$F(h,l) = -\sum_{v=h+1}^{l} D_{0,v}q^v \quad \text{when } 0 \leqq h < l \leqq b;$$

$$F(h,l) = 0 \quad \text{when } 0 \leqq h = l \leqq b;$$

$$F(h,l) = \sum_{v=l+1}^{h} D_{0,v}q^v \quad \text{when } 0 \leqq l < h \leqq b.$$

12. We are now ready to prove the theorem stated in the introduction. For this purpose we take $p=q$ and $a>b$ and suppose that G is primitive. Let s be an integer satisfying $0 \leqq s \leqq b-1$ and consider M_{p^{a-1},p^s}. Because of Lemma XIII, it is equal to at least one of the M_{i,jp^h} ($h=0, 1, \cdots,$ or b). Then because of Lemma XV, we have

$$\sum_{u=1}^{a-1}\sum_{v=0}^{s} D_{u,v}p^{u+v} = F(h, s). \tag{12.1}$$

Since $a+s \leqq a+b-1<a+b$, we also know that M_{p^a,p^s} must be equal to at least one of the M_{i,jp^h}, to $M_{i',j'p^{h'}}$, say. Then

$$\sum_{u=1}^{a}\sum_{v=0}^{s} D_{u,v}p^{u+v} = F(h', s). \tag{12.2}$$

(The values of h and h' are determined by the group G, but are unknown to us.) Subtracting (12.1) from (12.2), we find

$$\sum_{v=0}^{s} D_{a,v}p^{a+v} = F(h', s) - F(h, s). \tag{12.3}$$

Notice that for a fixed value of s the coefficient of $D_{0,v}$ in $F(h, s)$ if not zero is independent of h. Hence, if the values of h and h' are such that the coefficient of $D_{0,v}$ in neither $F(h, s)$ nor $F(h', s)$ is zero, the corresponding terms cancel in $F(h', s)-F(h, s)$.

We now take $s=0$ and try to show that $D_{a,0}=0$. For that purpose, it is sufficient to show that $|D_{a,0}p^a| < p^a$, for then $|D_{a,0}| < 1$, and since $D_{a,0}$ is an integer (Lemma XIV), it must be zero.

Since $p=q$,

$$\begin{aligned} F(0,0) &= 0, \\ F(1,0) &= D_{0,1}p, \\ F(2,0) &= D_{0,1}p + D_{0,2}p^2, \\ &\cdots\cdots\cdots\cdots\cdots \\ F(b,0) &= D_{0,1}p + D_{0,2}p^2 + \cdots + D_{0,b}p^b. \end{aligned}$$

If $b=1$, (12.3), Lemma XIV, and the condition $b<a$ used in turn give us

$$|D_{a,0}p^a| \leqq |D_{0,1}p| \leqq p < p^a.$$

There remains for consideration the case $b>1$.

If $h<b$ and $h'<b$,

$$\begin{aligned} |D_{a,0}p^a| &\leqq |D_{0,1}p| + |D_{0,2}p^2| + \cdots + |D_{0,b-1}p^{b-1}| \\ &\leqq 2p + 2p^2 + \cdots + 2p^{b-1} = 2(p^b - p)/(p-1) \\ &\leqq 2p^b - 2p < p^{b+1} \leqq p^a. \end{aligned}$$

If $h=h'=b$,

$$D_{a,0}p^a = 0, \qquad D_{a,0} = 0.$$

The remaining possibility is that $h=b$ and $h'<b$ or that $h'=b$ and $h<b$. Then if $D_{0,b}=0$,

$$\begin{aligned} |D_{a,0}p^a| &\leqq |D_{0,1}p| + |D_{0,2}p^2| + \cdots + |D_{0,b-1}p^{b-1}| + 0 \\ &\leqq 2p + 2p^2 + \cdots + 2p^{b-1} < p^a. \end{aligned}$$

If $D_{0,b}\neq 0$, $D_{0,b}=\pm 1$, by Lemma XIV. Then, since $D_{0,b-1}$ is an integer and $|D_{0,b-1}| \leqq 2$, $D_{0,b-1}$ is equal to $D_{0,b}$, $2D_{0,b}$, $-D_{0,b}$, $-2D_{0,b}$, or 0.

If $D_{0,b-1}=D_{0,b}$,

$$\begin{aligned} d_{0,b-1} - d_{0,b} - d_{1,b-1} + d_{1,b} &= d_{0,b} - d_{1,b}, \\ d_{0,b-1} - d_{1,b-1} &= 2(d_{0,b} - d_{1,b}), \\ d_{0,b-1} - d_{1,b-1} &\equiv 0 \bmod 2. \end{aligned}$$

But

$$|d_{0,b-1} - d_{1,b-1}| \leqq 1.$$

Therefore $d_{0,b-1}-d_{1,b-1}=0$, $d_{0,b}-d_{1,b}=0$, and $D_{0,b}=0$, contrary to our hypothesis.

The case $D_{0,b-1}=2D_{0,b}$ can also be shown to be impossible when $D_{0,b}\neq 0$, and in much the same way.

If $D_{0,b-1}=-D_{0,b}$, we consider separately the two possibilities $|h-h'|=1$

and $|h-h'|>1$. ($|h-h'|=0$ is impossible since, by hypothesis, one of the numbers h or h' is equal to b and the other is not.)

When $|h-h'|=1$, one of the two numbers h and h' is equal to b and the other to $b-1$. Then

$$|D_{a,0}p^a| = |D_{0,b}p^b| = p^b < p^a.$$

When $|h-h'|>1$, if $b=2$,

$$\begin{aligned}|D_{a,0}p^a| &= |D_{0,1}p + D_{0,2}p^2| \\ &= |D_{0,2}(-p+p^2)| \\ &= p^2 - p < p^2 < p^a.\end{aligned}$$

If $b>2$,

$$\begin{aligned}|D_{a,0}p^a| &\leqq |D_{0,1}p| + \cdots + |D_{0,b-2}p^{b-2}| + |D_{0,b-1}p^{b-1} + D_{0,b}p^b| \\ &\leqq 2p + \cdots + 2p^{b-2} + p^b - p^{b-1} \\ &= 2(p^{b-1} - p)/(p-1) + p^b - p^{b-1} \\ &\leqq 2p^{b-1} - 2p + p^b - p^{b-1} \\ &< p^b + p^{b-1} < 2p^b \leqq p^a.\end{aligned}$$

So that if $D_{0,b-1} = -D_{0,b}$, $|D_{a,0}p^a| < p^a$.

The case $D_{0,b-1} = -2D_{0,b}$ may be handled in much the same way. Again $|D_{a,0}p^a| < p^a$.

The fifth and last possibility was $D_{0,b-1}=0$. Then, if $b=2$,

$$|D_{a,0}p^a| \leqq 0 + |D_{0,2}p^2| = p^2 < p^a.$$

If $b>2$,

$$\begin{aligned}|D_{a,0}p^a| &\leqq |D_{0,1}p| + \cdots + |D_{0,b-2}p^{b-2}| + 0 + |D_{0,b}p^b| \\ &\leqq 2p + \cdots + 2p^{b-2} + p^b \\ &= 2(p^{b-1} - p)/(p-1) + p^b \\ &\leqq 2p^{b-1} - 2p + p^b < 2p^b \leqq p^a.\end{aligned}$$

So that in all cases $D_{a,0}=0$.

We now make use of this fact and (12.3) to calculate $D_{a,s}$ when $0<s<b$. We may write (12.3):

$$D_{a,s}p^{a+s} = F(h', s) - F(h, s) - \sum_{v=0}^{s-1} D_{a,v}p^{a+v},$$

so that

$$
\begin{aligned}
|D_{a,s}p^{a+s}| &\leqq |D_{0,1}p| + \cdots + |D_{0,b-1}p^{b-1}| + |D_{0,b}p^{b}| \\
&\quad + |D_{a,0}p^{a}| + |D_{a,1}p^{a+1}| + \cdots + |D_{a,s-1}p^{a+s-1}| \\
&\leqq 2p + \cdots + 2p^{b-1} + p^{b} + 0 + p^{a+1} + \cdots + p^{a+s-1} \\
&= 2(p^{b} - p)/(p-1) + p^{b} + (p^{a+s} - p^{a+1})/(p-1) \\
&\leqq 2p^{b} - 2p + p^{b} + p^{a+s} - p^{a+1} \\
&= p^{a+s} + (3p^{b} - p^{a+1}) - 2p < p^{a+s},
\end{aligned}
$$

$$D_{a,s} = 0.$$

Hence we have $D_{a,0}=D_{a,1}= \cdots =D_{a,b-1}=0$. But, by Lemma XVI, that is impossible when G is primitive. Therefore G is imprimitive when $p=q$ and $a\neq b$. That it is also compound follows from the last paragraph of the introduction.

13. The corresponding theorem when $p\neq q$ is covered by Schur's paper.* However, we are now in a position to prove it quite easily. And a proof based on the properties of group characteristics is not without interest, for the question of the relative advantages of using the theory of group characteristics (and the roots of unity which it involves) or of avoiding all irrational quantities as Schur did, is of importance to future investigations in the field. We therefore include the proof.

We may without loss of generality suppose that $p<q$. Assume that G is primitive. Then, by Lemma XIII, $M_{p,1}$ is equal to some M_{i,jq^h} with $h=0, 1, \cdots,$ or b.

If $h=0$, Lemma XV gives us

$$\sum_{u=1}^{1}\sum_{v=0}^{0} D_{u,v}p^{u}q^{v} = 0,$$

$$D_{1,0} = 0.$$

If $h>0$,

$$\sum_{u=1}^{1}\sum_{v=0}^{0} D_{u,v}p^{u}q^{v} = \sum_{v=1}^{h} D_{0,v}q^{v},$$

$$D_{1,0}p = D_{0,1}q + D_{0,2}q^{2} + \cdots + D_{0,h}q^{h},$$

$$D_{1,0} \equiv 0 \bmod q.$$

Since p and q are both primes and $p<q$, $q\geqq 3$. Now $|D_{1,0}| \leqq 2$ (by Lemma XIV). Therefore $D_{1,0}=0$.

Hence, regardless of h, $D_{1,0}=0$.

Suppose $D_{1,0}=D_{2,0}= \cdots =D_{k-1,0}=0$ and consider $M_{p^k,1}$ $(1<k\leqq a)$. By Lemma XIII, it is equal to some M_{i,jq^h} with $h=0, 1, \cdots,$ or b.

* *Zur Theorie der einfach transitiven Permutationsgruppen*, Sitzungsberichte der Preussischen Akademie der Wissenschaften, Physikalisch-mathematische Klasse, 1933, pp. 598–623.

If $h=0$, by Lemma XV,

$$\sum_{u=1}^{k}\sum_{v=0}^{0} D_{u,v}p^u q^v = 0,$$

$$D_{1,0}p + D_{2,0}p^2 + \cdots + D_{k-1,0}p^{k-1} + D_{k,0}p^k = 0,$$

$$D_{k,0}p^k = 0,$$

$$D_{k,0} = 0.$$

If $h>0$,

$$\sum_{u=1}^{k}\sum_{v=0}^{0} D_{u,v}p^u q^v = \sum_{v=1}^{h} D_{0,v}q^v,$$

$$D_{k,0}p^k = D_{0,1}q + D_{0,2}q^2 + \cdots + D_{0,h}q^h,$$

$$D_{k,0} \equiv 0 \bmod q,$$

$$D_{k,0} = 0.$$

So that, regardless of h, $D_{k,0}=0$.

Hence, by induction, $D_{1,0}=D_{2,0}=\cdots=D_{a,0}=0$.

Lemma XIII tells us that M_{p^k,q^l} is equal to at least one of the M_{ip^f,jq^g}. Then since $D_{1,0}=D_{2,0}=\cdots=D_{a,0}=0$, the second part of Lemma XV shows us that, regardless of the value of f,

$$\sum_{u=0}^{k}\sum_{v=1}^{l} D_{u,v}p^u q^v = 0.$$

By applying this equation using in turn the following sets of values for k and l:

$$\begin{gathered}(0,1),\quad (0,2),\quad \cdots,\ (0,b),\\ (1,1),\quad (1,2),\quad \cdots,\ (1,b),\\ \cdots\cdots\cdots\cdots\cdots,\\ (a-1,1),\ (a-1,2),\ \cdots,\ (a-1,b),\end{gathered}$$

we find $D_{u,v}=0$ ($u=0, 1, \cdots, a-1$; $v=1, 2, \cdots, b$). But $D_{0,b}=D_{1,b}=\cdots=D_{a-1,b}=0$ is impossible when G is primitive (Lemma XVI). Hence G is imprimitive when $p\neq q$.

14. Combining the results of §§12 and 13 with the remarks made in the introduction about imprimitive groups with transitive Abelian subgroups, we have the

THEOREM. *If p and q represent prime numbers, which may or may not be equal, and a and b positive integers, a simply transitive permutation group G of degree $n=p^aq^b$ which contains a transitive Abelian subgroup of degree n generated by two permutations, one of order p^a and one of order q^b, is imprimitive and has an invariant proper subgroup unless $p^a=q^b$.*

15. We are now ready to discuss Burnside's paper of 1921, which we mentioned in the introduction. It gives his proof of the theorem stated there when the Abelian subgroup can have two independent generators. The preliminary work is essentially that given in §1 of the present paper, the principal difference being that Burnside did not restrict the orders of the generators of the Abelian subgroup to powers of primes as has been done in this paper. The proof itself consists of two parts corresponding to the two cases in which (1) the order of at least one of the two generators is composite, and (2) the orders of both generators are primes.

In the first case, it will be sufficient to show that the work is incorrect when the orders of the generators are powers of primes. If, for example, p^a is composite, the statement from which Burnside concluded that G is compound says that P^p is represented by the identity of $\Gamma_{p^{a-1},0}$. However, that cannot in general be true, for Lemma I shows us that if it were true, G would be imprimitive. This is contradicted by the existence of exceptions, for example, the simply transitive primitive groups of degree k^2 and order $2(k!)^2$ mentioned in the introduction when $k=p^a$ with $a>1$. So that the first and most important part of the proof must contain a serious mistake.

The other case which Burnside takes up is that in which the two generators are of different prime orders. To quote from this paper (changing the notation to conform with that which we have been using): "If p and q are different primes, and Γ is that irreducible representation to which $\xi_{1,0}$ beongs, so that $\chi_{P^{a_u}Q^{b_u}}=(\mu/m)\sum_{u=1}^{m}\epsilon^{a_u}$, then

$$\frac{\mu}{m}\sum_{u=1}^{m}\epsilon^{a_u}=\sum_{v=1}^{\mu}\epsilon^{a_u\alpha_v}\eta^{b_u\beta_v}. \tag{15.1}$$

Unless each β_v is zero, in which case the group has a self-conjugate subgroup containing Q, this equation actually contains powers of η on the right. Hence when the indices of the powers of η are reduced (mod q) each power must occur with the same coefficient. This shews that μ must be a multiple of q, and $\cdots$." We cannot deny that μ is a multiple of q when at least one of the β_v is not equal to zero,* but we do deny that that conclusion can be drawn

* For the material of §§10–13 may be used to show that if the first two conditions of Lemma XII are not satisfied, either each transitive constituent of G_0 except $\Delta_{0,0}$ affects at least one $x_{i,s}$ $(i,\ p)=1$, or G is imprimitive and $p=q=2$. In the former case, we found that unless $p^a=q^b$, $D_{a,0}=\cdots=D_{a,b-1}=0$ or $D_{0,b}=\cdots=D_{a-1,b}=0$, both of which contradict the assumption that the first two conditions of Lemma XII are not satisfied. Therefore, except perhaps when $p=q=2$ or $p^a=q^b$, at least one of the first two conditions of Lemma XII is satisfied. We apply this fact and Lemma VII to $\Gamma_{p^{a-1},0}$ when $n=pq$ $(p\neq q)$. Either $d_{1,1}=d_{1,0}\neq d_{0,1}$ or $d_{1,1}=d_{0,1}\neq d_{1,0}$. But $d_{1,0}$ always equals one. Therefore if $d_{1,1}=d_{0,1}$, both equal zero, and each β_v equals zero. Hence if at least one of the β_v is not zero, $d_{1,1}=d_{1,0}=1$, $d_{0,1}=0$, and $\mu=c(q-1)+c=cq$.

directly from (15.1). For if we use the permutation PQ, (15.1) becomes

$$\frac{\mu}{m}\sum_{u=1}^{m}\epsilon^{\alpha_u}=\sum_{v=1}^{\mu}\epsilon^{\alpha_v\eta^{\beta_v}}. \tag{15.2}$$

It is possible to set up letters for the transitive constituent of G_0 which involves $x_{1,1}$ and variables for $\Gamma_{1,0}$ which satisfy the requirements imposed by (15.2) and for which μ is not a multiple of q. For example, if $p=3$ and $q=5$, we might propose $x_{1,1}$, $x_{2,1}$, $x_{0,1}$, $x_{1,2}$, $x_{2,2}$, $x_{1,3}$, $x_{2,3}$, $x_{1,4}$, $x_{2,4}$ as the letters of $\Delta_{1,1}$, and $\xi_{1,0}$, $\xi_{2,0}$, $\xi_{0,1}$, $\xi_{0,2}$, $\xi_{0,3}$, $\xi_{0,4}$ as the variables of $\Gamma_{1,0}$. Then $\mu=6$, $m=9$, and (15.2) becomes

$$\frac{6}{9}(\epsilon+\epsilon^2+1+\epsilon+\epsilon^2+\epsilon+\epsilon^2+\epsilon+\epsilon^2)=\epsilon+\epsilon^2+\eta+\eta^2+\eta^3+\eta^4,$$

and since ϵ now denotes a primitive cube root of unity and η a primitive fifth root of unity, this reduces to

$$-2=-2.$$

So that the proof of the second case is also incorrect.

The cases in which the Abelian subgroup cannot be generated by two generators Burnside dismisses with the statement, "It is clear that the same method of proof will apply, when the transitive Abelian subgroup has three or more independent generators." So the failure of his proof when the Abelian subgroup has two generators leaves the theorem unproved except in the cases dealt with elsewhere.

STANFORD UNIVERSITY,
PALO ALTO, CALIFORNIA

ON BOOLEAN FUNCTIONS OF MANY VARIABLES*

BY

J. C. C. McKINSEY

1. Introduction. I here treat of some facts regarding Boolean functions of n variables. The results are in part an extension to functions of n variables of theorems already known for functions of one variable, in part a discussion of new problems.

By a *function of n Boolean variables* I mean a rule whereby to each ordered set $x_1, \cdots, x_n$ of Boolean elements is assigned a Boolean element z. A function of Boolean variables is called a *Boolean function* if it can be expressed through a finite number of applications of the Boolean operations $+$, $\times$, and $'$. It is well known that every Boolean function $f(x_1, \cdots, x_n)$ can be expressed as $f(1, \cdots, 1)x_1 \cdots x_n + \cdots + f(0, \cdots, 0)x_1' \cdots x_n'$. By a *non-Boolean function*, I mean a function of Boolean variables which is not a Boolean function. I shall be concerned mainly, as already mentioned, with Boolean functions, but an occasional reference will be made to non-Boolean functions.

In §§2–6 I discuss "monotone" Boolean functions, which are analogous to the monotone functions of ordinary analysis. I establish conditions that functions be monotone non-decreasing and monotone non-increasing. In §§7–9 I discuss a more special kind of functions, namely, "additive" and "subtractive" functions. In §§10 and 11 I define and discuss an analogue, for Boolean functions of the "continuity" of classical analysis. I show here that all Boolean functions are "continuous" and suggest the importance of continuity with regard to non-Boolean functions. In §12 I discuss inverse functions, finding all the domains within which a function of n variables has a one-valued inverse.

Schmidt† obtained some interesting facts concerning Boolean functions of one variable. Schmidt's results are included in my Theorems 15, 17, 20, 24, and 30. Theorems 15, 17, 20, and 24 constitute a generalization to functions of n variables of his "Principal Theorem"‡ concerning Boolean functions of one variable. In §§5 and 8 I make some remarks on the ways in which this

* Presented to the Society, April 11, 1936; received by the editors February 28, 1936.

† Karl Schmidt, *The theory of functions of one Boolean variable*, these Transactions, vol. 23 (1922), pp. 212–222. All later references to Schmidt are to this paper.

‡ The theorem is as follows: *Every change of x, or part of x, in the increment region ab' produces the same change in ax+bx'; in the decrement region a'b it produces the opposite effect; and a change of x in any other region has no effect whatsoever on ax+bx', which always occupies ab and never enters a'b'.*

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

theorem fails to admit of generalization. Theorem 30 gives the domains within which a function of n variables has a one-valued inverse; this problem was only partially solved by Schmidt for functions of one variable.

Throughout the paper I use, besides the familiar Boolean operations, $+$, $\times$, and $'$, the operations $a \circ b$ and $a \Delta b$, which have been recently discussed by Bernstein.* These operations are defined as follows: $a \circ b = a'b + ab'$, $a \Delta b = ab + a'b'$. From their definitions it is easily seen that both operations enjoy the commutative and associative properties and that $a \Delta b = (a \circ b)'$. Other properties will be found in Bernstein's paper.

2. **Monotone non-decreasing functions.** One of the important general types of functions of a real variable is the set of monotone functions. A function of a real variable is *monotone non-decreasing* if it does not decrease when the argument increases, and *monotone non-increasing* if it does not increase when the argument increases. Thus, for example, the function $f(x) = \sin x$ is monotone non-decreasing in the interval $(0, \pi/2)$ and monotone non-increasing in the interval $(\pi/2, \pi)$. This classification rests upon the use of the relation $<$ as holding between real numbers, and hence cannot be applied directly to functions of Boolean variables. If we consider one of the most important interpretations of Boolean algebra, however, namely, the interpretation of the algebra as a calculus of classes, we are led to consider a function as monotone non-decreasing when it has the following characteristic: If the argument changes from one class to a more inclusive class, the function changes from one class to a more inclusive class. And analogously for monotone non-increasing functions. These notions are made more precise by the following definitions.

DEFINITION 1. A Boolean function $f(x_1, \cdots, x_n)$ is said to be *monotone non-decreasing with respect to x_1 in the domain (A, B)* if, for all α in (A, B),

$$f(x_1, \cdots, x_n) < f(x_1 + \alpha, x_2, \cdots, x_n)$$

independent of the choice of $x_1, \cdots, x_n$. If $f(x_1, \cdots, x_n)$ is monotone non-decreasing with respect to x_1 in the domain $(0, 1)$, we say that $f(x_1, \cdots, x_n)$ is *monotone non-decreasing with respect to x_1 everywhere.*

DEFINITION 2. A Boolean function $f(x_1, \cdots, x_n)$ is said to be *monotone non-increasing with respect to x_1 in the domain (A, B)* if, for all α in (A, B),

$$f(x_1 + \alpha, x_2, \cdots, x_n) < f(x_1, \cdots, x_n)$$

independent of the choice of $x_1, \cdots, x_n$. If $f(x_1, \cdots, x_n)$ is monotone non-

* B. A. Bernstein, *Postulates for Boolean algebra involving the operation of complete disjunction*, to appear in the Annals of Mathematics. I take this opportunity of acknowledging my indebtedness to Professor Bernstein for his valuable suggestions regarding the present paper.

increasing with respect to x_1 in the domain (0, 1), we say that $f(x_1, \cdots, x_n)$ is *monotone non-increasing with respect to x_1 everywhere.*

Thus, for example, the Boolean function of one variable $f(x) = ax$ is monotone non-decreasing everywhere; for,

$$f(x) = ax < ax + a\alpha = a(x + \alpha) = f(x + \alpha)$$

for any x. Similarly the function $f(x) = ax'$ is monotone non-increasing everywhere; for,

$$f(x + \alpha) = a(x + \alpha)' = ax'\alpha' < ax' = f(x)$$

for any x.

I now find a condition that a function of one variable be monotone non-decreasing in a domain.

THEOREM 1. *A necessary and sufficient condition that a given function $f(x)$ of one variable be monotone non-decreasing in the domain (A, B) is that*

$$B < f(1) + f'(0).$$

Proof. To say that $f(x)$ is monotone non-decreasing in (A, B) is, by definition, equivalent to saying that for all α in (A, B) we have

$$f(x) < f(x + \alpha). \tag{1}$$

But (1) is equivalent to

$$f(x)f'(x + \alpha) = 0, \tag{2}$$

which is equivalent to

$$f(0)f'(1)\alpha x' = 0. \tag{3}$$

Since (3) is an identity in x, it is equivalent to

$$f(0)f'(1)\alpha = 0, \tag{4}$$

or to

$$\alpha < f(1) + f'(0). \tag{5}$$

If (5) holds for all α in (A, B), it holds in particular for $\alpha = B$, since B is in (A, B); moreover, if (5) holds for B, it holds for all α in (A, B), since if α is in (A, B) we have $\alpha < B$. Hence to say that (5) holds for all α in (A, B) is equivalent to saying that

$$B < f(1) + f'(0), \tag{6}$$

as was to be shown.

I next generalize Theorem 1 to functions of n variables.

THEOREM 2. *A necessary and sufficient condition that a given function* $f(x_1, \cdots, x_n)$ *be monotone non-decreasing with respect to* x_1 *in the domain* (A, B) *is that*

$$B < \prod_{\alpha_i=0,1} [f(1, \alpha_2, \cdots, \alpha_n) + f'(0, \alpha_2, \cdots, \alpha_n)].$$

Proof. I prove the theorem by mathematical induction on n.

For a function of one variable, we have Theorem 1.

Suppose, now, that the theorem is true for $n=k$. Let $f(x_1, \cdots, x_{k+1})$ be any function of $k+1$ variables. For any fixed value, z, of x_{k+1}, $f(x_1, \cdots, x_k, z)$ is a function of k variables. Then, since by hypothesis the theorem holds for a function of k variables, we have: A necessary and sufficient condition that $f(x_1, \cdots, x_k, z)$ be monotone non-decreasing with respect to x_1 for the domain (A, B) is that

$$B < \prod_{\alpha_i=0,1} [f(1, \alpha_2, \cdots, \alpha_k, z) + f'(0, \alpha_2, \cdots, \alpha_k, z)]. \tag{1}$$

Since (1) is the required necessary and sufficient condition for each z, we may regard z as a variable in the condition. Thus, a necessary and sufficient condition that $f(x_1, \cdots, x_n)$ be monotone non-decreasing in (A, B) is that (1) hold for all z. But to say that (1) holds for all z is equivalent to saying

$$\begin{cases} B < \prod_{\alpha_i=0,1} [f(1, \alpha_2, \cdots, \alpha_k, 1) + f'(0, \alpha_2, \cdots, \alpha_k, 1)] \\ B < \prod_{\alpha_i=0,1} [f(1, \alpha_2, \cdots, \alpha_k, 0) + f'(0, \alpha_2, \cdots, \alpha_k, 0)]. \end{cases} \tag{2}$$

And conditions (2) are equivalent to the condition specified in the theorem. Hence the theorem is proved.

THEOREM 3. *A necessary and sufficient condition that a function* $f(x_1, \cdots, x_n)$ *be monotone non-decreasing with respect to* x_1 *everywhere is that*

$$f(0, \alpha_2, \cdots, \alpha_n) < f(1, \alpha_2, \cdots, \alpha_n) \qquad [\alpha_i = 0, 1].$$

Proof. For, to say $f(x_1, \cdots, x_n)$ is monotone non-decreasing with respect to x_1 everywhere is, by definition, to say that $f(x_1, \cdots, x_n)$ is monotone non-decreasing with respect to x_1 in the domain $(0, 1)$. This, by Theorem 2, is equivalent to

$$\prod_{\alpha_i=0,1} [f(1, \alpha_2, \cdots, \alpha_n) + f'(0, \alpha_2, \cdots, \alpha_n)] = 1,$$

or to

$$f(1, \alpha_2, \cdots, \alpha_n) + f'(0, \alpha_2, \cdots, \alpha_n) = 1 \qquad [\alpha_i = 0, 1].$$

The last equations are equivalent to the given inclusions.

THEOREM 4. *A necessary and sufficient condition that a function* $f(x_1, \cdots, x_n)$ *be monotone non-decreasing with respect to* x_1 *everywhere is that* $f(x_1, \cdots, x_n)$ *be expressible in a form which does not involve* x_1' *(but which may involve* x_1*).*

Proof. If $f(x_1, \cdots, x_n)$ is monotone non-decreasing with respect to x_1 everywhere, then, by the preceding theorem, we have

$$f(0, \alpha_2, \cdots, \alpha_n) < f(1, \alpha_2, \cdots, \alpha_n) \qquad [\alpha_i = 0, 1].$$

Hence

$$f(0, \alpha_2, \cdots, \alpha_n) = f(0, \alpha_2, \cdots, \alpha_n) f(1, \alpha_2, \cdots, \alpha_n) \qquad [\alpha_i = 0, 1].$$

Thus

$$\begin{aligned} f(x_1, \cdots, x_n) &= f(1, \cdots, 1)x_1 \cdots x_n + \cdots + f(1, 0, \cdots, 0)x_1x_2' \cdots x_n' \\ &\quad + f(0, 1, \cdots, 1)f(1, \cdots, 1)x_1' x_2 \cdots x_n + \cdots \\ &\quad + f(0, \cdots, 0)f(1, 0, \cdots, 0)x_1' \cdots x_n' \\ &= f(1, \cdots, 1)[x_1 + x_1' f(0, 1, \cdots, 1)]x_2 \cdots x_n + \cdots \\ &\quad + f(1, 0, \cdots, 0)[x_1 + x_1' f(0, \cdots, 0)]x_2' \cdots x_n' \\ &= f(1, \cdots, 1)[x_1 + f(0, 1, \cdots, 1)]x_2 \cdots x_n + \cdots \\ &\quad + f(1, 0, \cdots, 0)[x_1 + f(0, \cdots, 0)]x_2' \cdots x_n', \end{aligned}$$

which is a form not involving x_1'. Hence, the condition is necessary.

Suppose, on the other hand, that $f(x_1, \cdots, x_n)$ can be expressed in a form not involving x_1'. The most general such form is

$$f(x_1, \cdots, x_n) = x_1 g(x_2, \cdots, x_n) + h(x_2, \cdots, x_n).$$

Then we have

$$\begin{aligned} f(0, \alpha_2, \cdots, \alpha_n) &= h(\alpha_2, \cdots, \alpha_n), \\ f(1, \alpha_2, \cdots, \alpha_n) &= g(\alpha_2, \cdots, \alpha_n) + h(\alpha_2, \cdots, \alpha_n). \end{aligned}$$

Therefore,

$$f(0, \alpha_2, \cdots, \alpha_n) < f(1, \alpha_2, \cdots, \alpha_n) \qquad [\alpha_i = 0, 1],$$

and hence, again by the preceding theorem, $f(x_1, \cdots, x_n)$ is monotone non-decreasing with respect to x_1 everywhere. Hence, the condition is sufficient.

THEOREM 5. *A necessary and sufficient condition that* $f(x_1, \cdots, x_n)$ *be monotone non-decreasing with respect to* x_1 *everywhere is that, for every* u *and* v,

$$f(u + v, x_2, \cdots, x_n) = f(u, x_2, \cdots, x_n) + f(v, x_2, \cdots, x_n).$$

Proof. To say that

$$(1) \qquad f(u+v, x_2, \cdots, x_n) = f(u, x_2, \cdots, x_n) + f(v, x_2, \cdots, x_n)$$

holds for all u and v, is equivalent to saying it holds for $u, v=0, 1$. But (1) becomes an identity when we substitute $u=v=1$, and likewise when we substitute $u=v=0$. Furthermore, we obtain the same condition when we put $u=1, v=0$ as when we put $u=0, v=1$. Hence to say that (1) holds for all u and v is equivalent to saying that

$$(2) \qquad f(1, x_2, \cdots, x_n) = f(1, x_2, \cdots, x_n) + f(0, x_2, \cdots, x_n),$$

which is in turn equivalent to

$$(3) \qquad f(0, x_2, \cdots, x_n) < f(1, x_2, \cdots, x_n).$$

Since (3) is an identity in $x_2, \cdots, x_n$ it is the same as the condition, given in Theorem 3, that $f(x_1, \cdots, x_n)$ be monotone non-decreasing with respect to x_1 everywhere. Hence, (1) is a necessary and sufficient condition that $f(x_1, \cdots, x_n)$ be monotone non-decreasing with respect to x_1 everywhere.

The following theorem is proved in a similar way.

THEOREM 6. *A necessary and sufficient condition that $f(x_1, \cdots, x_n)$ be monotone non-decreasing with respect to x_1 everywhere is that, for every u and v*

$$f(uv, x_2, \cdots, x_n) = f(u, x_2, \cdots, x_n) f(v, x_2, \cdots, x_n).$$

The following theorem gives a condition for monotone non-decreasing domains which might have been taken as the definition. Definition 1 was selected only because its content is slightly more obvious intuitively.

THEOREM 7. *A necessary and sufficient condition that a function $f(x_1, \cdots, x_n)$ be monotone non-decreasing with respect to x_1 in the domain (A, B) is that, for all α in (A, B), we have*

$$f(x_1 \alpha', x_2, \cdots, x_n) < f(x_1, x_2, \cdots, x_n)$$

independent of the choice of $x_1, \cdots, x_n$.

Proof. Let $f(x)$ be a function of one variable which satisfies

$$(1) \qquad f(x\alpha') < f(x)$$

identically for all α in (A, B). Condition (1) reduces to

$$(2) \qquad f'(1) f(0) \alpha x = 0.$$

Since (2) is an identity in x, it is equivalent to

$$(3) \qquad f'(1) f(0) \alpha = 0,$$

which is equivalent to

$$\alpha < f(1) + f'(0). \tag{4}$$

Thus,

$$B < f(1) + f'(0). \tag{5}$$

It may now be shown, by mathematical induction on n, that $f(x_1, \cdots, x_n)$ satisfies

$$f(x_1\, \alpha', x_2, \cdots, x_n) < f(x_1, \cdots, x_n)$$

identically for all α in (A, B) if and only if

$$B < \prod_{\alpha_i=0,1} [f(1, \alpha_2, \cdots, \alpha_n) + f'(0, \alpha_2, \cdots, \alpha_n)].$$

But this is the condition, found in Theorem 2, that $f(x_1, \cdots, x_n)$ be monotone non-decreasing in (A, B).

3. **Monotone non-increasing functions.** The proofs of the following seven theorems are omitted, since they are closely analogous to the proofs of the corresponding theorems on monotone non-decreasing functions.

THEOREM 8. *A necessary and sufficient condition that a given function $f(x)$ of one variable be monotone non-increasing in the domain (A, B) is that*

$$B < f'(1) + f(0).$$

THEOREM 9. *A necessary and sufficient condition that a given function $f(x_1, \cdots, x_n)$ be monotone non-increasing with respect to x_1 in the domain (A, B) is that*

$$B < \prod_{\alpha_i=0,1} [f'(1, \alpha_2, \cdots, \alpha_n) + f(0, \alpha_2, \cdots, \alpha_n)].$$

THEOREM 10. *A necessary and sufficient condition that a function $f(x_1, \cdots, x_n)$ be monotone non-increasing with respect to x_1 everywhere is that*

$$f(1, \alpha_2, \cdots, \alpha_n) < f(0, \alpha_2, \cdots, \alpha_n) \qquad [\alpha_i = 0, 1].$$

THEOREM 11. *A necessary and sufficient condition that a function $f(x_1, \cdots, x_n)$ be monotone non-increasing with respect to x_1 everywhere is that $f(x_1, \cdots, x_n)$ be expressible in a form which does not involve x_1 (but which may involve x_1').*

THEOREM 12. *A necessary and sufficient condition that $f(x_1, \cdots, x_n)$ be monotone non-increasing with respect to x_1 everywhere is that, for every u and v,*

$$f(u + v, x_2, \cdots, x_n) = f(u, x_2, \cdots, x_n) f(v, x_2, \cdots, x_n).$$

THEOREM 13. *A necessary and sufficient condition that* $f(x_1, \cdots, x_n)$ *be monotone non-increasing with respect to* x_1 *everywhere is that, for every* u *and* v,

$$f(uv, x_2, \cdots, x_n) = f(u, x_2, \cdots, x_n) + f(v, x_2, \cdots, x_n).$$

THEOREM 14. *A necessary and sufficient condition that a function* $f(x_1, \cdots, x_n)$ *be monotone non-increasing with respect to* x_1 *in the domain* (A, B) *is that, for all* α *in* (A, B), *we have*

$$f(x_1, x_2, \cdots, x_n) < f(x_1 \alpha', x_2, \cdots, x_n)$$

independent of the choice of $x_1, \cdots, x_n$.

4. **Functions both monotone non-decreasing and monotone non-increasing. Constant functions.** We prove the following theorem.

THEOREM 15. *A necessary and sufficient condition that a function* $f(x_1, \cdots, x_n)$ *be both monotone non-decreasing and monotone non-increasing with respect to* x_1 *in the domain* (A, B) *is that*

$$B < \prod_{\alpha_i=0,1} [f(1, \alpha_2, \cdots, \alpha_n) \Delta f(0, \alpha_2, \cdots, \alpha_n)].$$

Proof. To say that $f(x_1, \cdots, x_n)$ is both monotone non-decreasing and monotone non-increasing with respect to x_1 in (A, B) is, by Theorems 2 and 9, equivalent to saying that

$$(1) \qquad \begin{cases} B < \prod_{\alpha_i=0,1} [f(1, \alpha_2, \cdots, \alpha_n) + f'(0, \alpha_2, \cdots, \alpha_n)], \\ B < \prod_{\alpha_i=0,1} [f'(1, \alpha_2, \cdots, \alpha_n) + f(0, \alpha_2, \cdots, \alpha_n)]. \end{cases}$$

Conditions (1) are equivalent to the single condition

$$(2) \qquad \begin{aligned} B < &\prod_{\alpha_i=0,1} [f(1, \alpha_2, \cdots, \alpha_n) + f'(0, \alpha_2, \cdots, \alpha_n)] \\ &\cdot \prod_{\alpha_i=0,1} [f'(1, \alpha_2, \cdots, \alpha_n) + f(0, \alpha_2, \cdots, \alpha_n)], \end{aligned}$$

or

$$(3) \qquad \begin{aligned} B < &\prod_{\alpha_i=0,1} [f(1, \alpha_2, \cdots, \alpha_n) + f'(0, \alpha_2, \cdots, \alpha_n)] \\ &\cdot [f'(1, \alpha_2, \cdots, \alpha_n) + f(0, \alpha_2, \cdots, \alpha_n)], \end{aligned}$$

or, finally,

$$(4) \qquad B < \prod_{\alpha_i=0,1} [f(1, \alpha_2, \cdots, \alpha_n) \Delta f(0, \alpha_2, \cdots, \alpha_n)].$$

THEOREM 16. *A necessary and sufficient condition that* $f(x_1, \cdots, x_n)$ *be both monotone non-decreasing and monotone non-increasing with respect to* x_1 *everywhere is that*

$$f(1, \alpha_2, \cdots, \alpha_n) = f(0, \alpha_2, \cdots, \alpha_n) \qquad [\alpha_i = 0, 1],$$

hence, that $f(x_1, \cdots, x_n)$ *be expressible in a form which involves neither* x_1 *nor* x_1'.

Proof. The theorem is an immediate consequence of Theorem 15.

By saying that $f(x_1, \cdots, x_n)$ is *constant with respect to* x_1 *in the domain* (A, B), I mean that $f(x_1, \cdots, x_n)$ is unchanged when x_1 assumes or loses any increment from (A, B).

THEOREM 17. *A necessary and sufficient condition that* $f(x_1, \cdots, x_n)$ *be constant with respect to* x_1 *in a domain* (A, B) *is that* $f(x_1, \cdots, x_n)$ *be both monotone non-decreasing and monotone non-increasing with respect to* x_1 *in the domain* (A, B).

Proof. To say that $f(x_1, \cdots, x_n)$ is constant with respect to x_1 in the domain (A, B) is to say that for all α in (A, B)

$$\begin{aligned} f(x_1 + \alpha, x_2, \cdots, x_n) &= f(x_1, \cdots, x_n), \\ f(x_1\alpha', x_2, \cdots, x_n) &= f(x_1, \cdots, x_n). \end{aligned} \tag{1}$$

Conditions (1) are equivalent to the four conditions

$$\begin{cases} f(x_1 + \alpha, x_2, \cdots, x_n) < f(x_1, \cdots, x_n), \\ f(x_1 + \alpha, x_2, \cdots, x_n) > f(x_1, \cdots, x_n), \\ f(x_1\alpha', x_2, \cdots, x_n) < f(x_1, \cdots, x_n), \\ f(x_1\alpha', x_2, \cdots, x_n) > f(x_1, \cdots, x_n). \end{cases} \tag{2}$$

By Theorems 7 and 14, respectively, the third and fourth conditions of (2) are consequences of the first and second. Hence conditions (1) are equivalent to

$$\begin{aligned} f(x_1 + \alpha, x_2, \cdots, x_n) &< f(x_1, \cdots, x_n), \\ f(x_1 + \alpha, x_2, \cdots, x_n) &> f(x_1, \cdots, x_n). \end{aligned} \tag{3}$$

Hence the theorem is established.

THEOREM 18. *A necessary and sufficient condition that* $f(x_1, \cdots, x_n)$ *be constant with respect to* x_1 *everywhere is that*

$$f(1, \alpha_2, \cdots, \alpha_n) = f(0, \alpha_2, \cdots, \alpha_n) \qquad [\alpha_i = 0, 1],$$

hence, that $f(x_1, \cdots, x_n)$ *be expressible in a form which involves neither* x_1 *nor* x_1'.

Proof. The theorem is an immediate consequence of Theorems 16 and 17.

5. Remarks on the variation of functions. It is of some interest to notice that the theorems given so far are sufficient to give a complete, though rough, characteriziation of the variation of a function of one variable. When we are given the function $f(x)=ax+bx'$, then we may consider the universe as falling into the four regions: ab, ab', $a'b$, $a'b'$. From Theorems 15, 17, 1, 7, 8, and 14 we now see that $f(x)$ is affected as follows when x assumes or loses increments from these regions:

region	if x assumes an increment	if x loses an increment
ab	$f(x)$ is unchanged	$f(x)$ is unchanged
$a'b'$	$f(x)$ is unchanged	$f(x)$ is unchanged
ab'	$f(x)$ assumes an increment*	$f(x)$ loses an increment
$a'b$	$f(x)$ loses an increment	$f(x)$ assumes an increment

For functions of more than one variable, however, the theorems are not sufficient to allow such a complete characterization. For a function of two variables, for example, $f(x, y)=axy+bxy'+cx'y+dx'y'$, we may consider the universe as falling into sixteen regions:

$abcd$,	$abc'd$,	$a'b'cd$,	$a'bcd'$,
$ab'cd'$,	$abcd'$,	$a'b'cd'$,	$ab'c'd$.
$a'bc'd$,	$abc'd'$,	$a'b'c'd$,	
$a'b'c'd'$,	$ab'c'd'$,	$a'bcd$,	
	$a'bc'd'$,	$ab'cd'$,	

From Theorems 15 and 17, it is seen that the four regions in the first column are such that, when x assumes or loses an increment from one of them, $f(x, y)$ is unchanged; they are thus like the regions ab and $a'b'$ for functions of one variable. From Theorems 1 and 7, we see that the five regions in the second column are monotone non-decreasing and not constant; they are thus like region ab' for one variable. And from Theorems 8 and 14, we see that the five regions in the third column are monotone non-increasing and not constant, and are thus like region $a'b$ for one variable. Of the two regions in the last column, however, we know only that $f(x, y)$ is neither monotone non-decreasing nor monotone non-increasing for changes of x in them.

6. Duality considerations for monotone functions. By Theorem 1, to say that $f(x)=ax+bx'$ is monotone non-decreasing in (A, B) is equivalent to saying

$$A < B < a + b'. \tag{1}$$

* The increment assumed, or lost, by $f(x)$ may, however, in any one of the cases, be vanishing.

The dual of (1) is, if we represent the dual of A by A_1, etc.,

$$A_1 > B_1 > a_1 b_1' \tag{2}$$

or

$$A_1' < B_1' < a_1' + b_1. \tag{3}$$

By Theorem 8, to say that $g(x) = a_1x + b_1x'$ is monotone non-increasing in (A_1', B_1') is equivalent to saying

$$A_1' < B_1' < a_1' + b_1. \tag{4}$$

But (4) is the same as (3). Hence, the dual of "$ax+bx'$ is monotone non-decreasing in (A, B)" is "a_1x+b_1x' is monotone non-increasing in (A_1', B_1')."

Similar considerations apply to functions of n variables.

It is now seen that Theorem 8 is the dual of Theorem 1. For, writing $ax+bx'$ for $f(x)$, a for $f(1)$ and b for $f(0)$ in Theorem 1, and applying the rule, we obtain the following proposition: A necessary and sufficient condition that a_1x+b_1x' be monotone non-increasing in (A_1', B_1') is that $B_1 > a_1b_1'$. Substituting a_1 for a, b_1 for b, A_1' for A, and B_1' for B, we obtain: A necessary and sufficient condition that $ax+bx'$ be monotone non-increasing in (A, B) is that $B' > ab'$. But $B' > ab'$ is equivalent to $B < a'+b$, so we have Theorem 8.

In a similar way, it can be shown that Theorems 9, 10, 12, 13, and 14 are the respective duals of Theorems 2, 3, 5, 6, and 7. Theorems 15, 16, 17, and 18 are self-dual.

7. **Additive and subtractive functions.** Schmidt has established the following fact concerning a Boolean function of one variable, $f(x) = ax+bx'$: If $\alpha < ab'$, then $f(x+\alpha) = f(x)+\alpha$ and $f(x\alpha') = f(x)\alpha'$; and if $\alpha < a'b$, then $f(x+\alpha) = f(x)\alpha'$ and $f(x\alpha') = f(x)+\alpha$.

I propose in this section to generalize this result to functions of n variables. To this end I first lay down the following definitions.

DEFINITION 3. A function $f(x_1, \cdots, x_n)$ is said to be *additive with respect to x_1 in the domain* (A, B) if, for all α in (A, B), we have

$$f(x_1 + \alpha, x_2, \cdots, x_n) = f(x_1, \cdots, x_n) + \alpha,$$
$$f(x_1\alpha', x_2, \cdots, x_n) = f(x_1, \cdots, x_n)\alpha'$$

independent of the choice of $x_1, \cdots, x_n$. If $f(x_1, \cdots, x_n)$ is additive with respect to x_1 in the domain $(0, 1)$, we say that $f(x_1, \cdots, x_n)$ is *additive with respect to x_1 everywhere.*

DEFINITION 4. A function $f(x_1, \cdots, x_n)$ is said to be *subtractive with respect to x_1 in the domain* (A, B) if, for all α in (A, B), we have

$$f(x_1 + \alpha, x_2, \cdots, x_n) = f(x_1, \cdots, x_n)\alpha',$$
$$f(x_1\alpha', x_2, \cdots, x_n) = f(x_1, \cdots, x_n) + \alpha$$

independent of the choice of $x_1, \cdots, x_n$. If $f(x_1, \cdots, x_n)$ is subtractive with respect to x_1 in the domain $(0, 1)$, we say that $f(x_1, \cdots, x_n)$ is subtractive with respect to x_1 everywhere.

The following theorem includes the result, mentioned above, due to Schmidt, but is somewhat more complete, in that it shows that the given condition is necessary as well as sufficient.

THEOREM 19. *A necessary and sufficient condition that a function of one variable $f(x)$ be additive in (A, B) is that $B < f(1)f'(0)$.*

Proof. The condition

$$f(x + \alpha) = f(x) + \alpha \tag{1}$$

reduces to

$$f'(1)\alpha x + f'(1)\alpha x' = 0, \tag{2}$$

or

$$f'(1)\alpha = 0. \tag{3}$$

The condition

$$f(x\alpha') = f(x)\alpha' \tag{4}$$

reduces to

$$f(0)\alpha x + f(0)\alpha x' = 0, \tag{5}$$

or

$$f(0)\alpha = 0. \tag{6}$$

Equations (3) and (6) are together equivalent to

$$\alpha[f'(1) + f(0)] = 0, \tag{7}$$

or

$$\alpha < f(1)f'(0). \tag{8}$$

By an argument like that used in the proof of Theorem 1, it is now seen that to say that (8) holds for all α in (A, B) is equivalent to saying that

$$B < f(1)f'(0). \tag{9}$$

In the above proof, it will be noticed that the condition that (2) hold for *all* x is the same as the condition that it have a solution for *some* x; and similarly, the condition that (5) hold for *all* x is the same as that it have a solution for *some* x. Hence, we could weaken Definition 3 (as applying to functions of one variable) to say that $f(x)$ is *additive in* (A, B) *if for all* α *in* (A, B) *there is at least one* x *such that*

$$f(x + \alpha) = f(x) + \alpha$$

and at least one x *such that*

$$f(x\alpha') = f(x)\alpha'.$$

A similar situation will be found to obtain with respect to Definition 4.

I next generalize Theorem 19 to functions of n variables.

THEOREM 20. *A necessary and sufficient condition that* $f(x_1, \cdots, x_n)$ *be additive with respect to* x_1 *in* (A, B) *is that*

$$B < \prod_{\alpha_i=0,1} f(1, \alpha_2, \cdots, \alpha_n) f'(0, \alpha_2, \cdots, \alpha_n).$$

Proof. For $n=1$ the theorem reduces to Theorem 19.

Suppose the theorem true for $n=k$, and let $f(x_1, \cdots, x_{k+1})$ be a function of $k+1$ variables. Then, for any given value of x_{k+1}, as y, $f(x_1, \cdots, x_k, y)$ is a function of k variables; hence, by hypothesis, a necessary and sufficient condition that $f(x_1, \cdots, x_k, y)$ be additive with respect to x_1 in (A, B) is that

$$B < \prod_{\alpha_i=0,1} f(1, \alpha_2, \cdots, \alpha_k, y) f'(0, \alpha_2, \cdots, \alpha_k, y).$$

Hence, a necessary and sufficient condition that $f(x_1, \cdots, x_{k+1})$ be additive is that (1) hold for all y. But to say that (1) holds for all y is equivalent to saying that it holds for $y=1$ and $y=0$, hence to

$$B < \prod_{\alpha_i=0,1} f(1, \alpha_2, \cdots, \alpha_k, 1) f'(0, \alpha_2, \cdots, \alpha_k, 1),$$

$$B < \prod_{\alpha_i=0,1} f(1, \alpha_2, \cdots, \alpha_k, 0) f'(0, \alpha_2, \cdots, \alpha_k, 0).$$

These two inclusions are equivalent to the single inclusion

$$B < \prod_{\alpha_i=0,1} f(1, \alpha_2, \cdots, \alpha_{k+1}) f'(0, \alpha_2, \cdots, \alpha_{k+1}).$$

Hence, the theorem holds for $n=k+1$ if it holds for $n=k$.

THEOREM 21. *A necessary and sufficient condition that* $f(x_1, \cdots, x_n)$ *be additive with respect to* x_1 *everywhere is that*

$$f(1, \alpha_2, \cdots, \alpha_n) = 1, \qquad f(0, \alpha_2, \cdots, \alpha_n) = 0 \qquad [\alpha_i = 0, 1].$$

The theorem follows from Theorem 20.

THEOREM 22. *The only function that is additive with respect to x_1 everywhere is*

$$f(x_1, \cdots, x_n) = x_1.$$

Proof. Substituting the values given by Theorem 21 into the identity

$$f(x_1, \cdots, x_n) = f(1, \cdots, 1)x_1 \cdots x_n + \cdots + f(0, \cdots, 0)x_1' \cdots x_n',$$

we find

$$\begin{aligned} f(x_1, \cdots, x_n) &= x_1 \cdots x_n + \cdots + x_1 x_2' \cdots x_n' \\ &= x_1(x_2 \cdots x_n + \cdots + x_2' \cdots x_n') \\ &= x_1. \end{aligned}$$

The following theorems on subtractive functions are stated without proof, since the proofs are analogous to the corresponding proofs for additive functions.

THEOREM 23. *A necessary and sufficient condition that a function of one variable $f(x)$ be subtractive in (A, B) is that*

$$B < f'(1)f(0).$$

THEOREM 24. *A necessary and sufficient condition that $f(x_1, \cdots, x_n)$ be subtractive with respect to x_1 in (A, B) is that*

$$B < \prod_{\alpha_i=0,1} f'(1, \alpha_2, \cdots, \alpha_n)f(0, \alpha_2, \cdots, \alpha_n).$$

THEOREM 25. *A necessary and sufficient condition that $f(x_1, \cdots, x_n)$ be subtractive with respect to x_1 everywhere is that*

$$f(1, \alpha_2, \cdots, \alpha_n) = 0, \qquad f(0, \alpha_2, \cdots, \alpha_n) = 1 \qquad [\alpha_i = 0, 1].$$

THEOREM 26. *The only function that is subtractive with respect to x_1 everywhere is*

$$f(x_1, \cdots, x_n) = x_1'.$$

8. **Remarks on additive and subtractive functions.** It will doubtless have been noticed that, for a function of one variable, the region, ab', in which $f(x)$ is monotone non-decreasing and not constant, coincides with the region in which $f(x)$ is additive. Such is not the case, however, for functions of two or more variables. Thus, for example, for a function of two variables, $f(x, y) = axy + bxy' + cx'y + dx'y'$, $f(x, y)$ is monotone non-decreasing with respect to x and not constant in the five regions $abc'd$, $abcd'$, $abc'd'$, $ab'c'd'$,

$a'bc'd$, while the only region in which $f(x, y)$ is additive with respect to x is $abc'd'$. It is also of some interest to observe that a function of n variables is never additive in the same domain with respect to two or more variables.

Analogous remarks apply to subtractive domains.

9. **Duality considerations.** By an argument similar to that used in §6, it can be shown that the dual of "$ax+bx'$ is additive in (A, B)" is "a_1x+b_1x' is subtractive in (A_1', B_1')." Thus, Theorems 23, 24, 25, and 26 are the duals, respectively, of Theorems 19, 20, 21, and 22.

10. **Continuity.** In ordinary analysis, we say that a function $f(x)$ is *uniformly continuous* if for every $\epsilon>0$ there is a $\delta>0$ such that

$$\text{if} \quad \{\,|\,x-y\,|<\delta\} \quad \text{then} \quad \{\,|\,f(x)-f(y)\,|<\epsilon\}.$$

It is of some interest to try to extend the notion of "uniform continuity" to Boolean algebra.

To make such an extension it is first necessary to find in Boolean algebra some analogue of the operation $|x-y|$. Probably the best such analogue* is $x \circ y$. From the usual geometrical representation of $x \circ y$ it is seen that $x \circ y$ is the total amount by which x and y fail to coincide; if $x \circ y=0$, then $x=y$.

On the basis of this analogy, we may lay down the following definition.

DEFINITION 5. A function $f(x)$ of a Boolean variable is said to be *uniformly continuous* if, for every Boolean element $\epsilon\neq 0$, there exists an element $\delta\neq 0$ such that

$$\text{if} \quad \{x \circ y<\delta\} \quad \text{then} \quad \{f(x) \circ f(y)<\epsilon\}.$$

The definition can be extended to functions of n variables as follows.

DEFINITION 6. A function $f(x_1, \cdots, x_n)$ of n Boolean variables is said to be *uniformly continuous* if, for every Boolean element $\epsilon\neq 0$, there exists an ordered set of Boolean elements $\delta_1, \cdots, \delta_n$, with $\delta_i\neq 0$, such that

$$\text{if } \{(x_1 \circ y_1<\delta_1), \cdots, (x_n \circ y_n<\delta_n)\} \text{ then } \{f(x_1, \cdots, x_n) \circ f(y_1, \cdots, y_n)<\epsilon\}.$$

The concept just defined, however, despite the fact that it is analogous to one of the dominant concepts of ordinary analysis, is not a very important one in connection with Boolean functions, for, as shown in the following theorems, every Boolean function is uniformly continuous.

THEOREM 27. *If $f(x_1, \cdots, x_n)$ is any Boolean function, then*

$$f(x_1, \cdots, x_n) \circ f(y_1, \cdots, y_n)<(x_1 \circ y_1)+\cdots+(x_n \circ y_n).$$

* For a treatment of $x \circ y$ as an analogue of $|x-y|$ see P. J. Daniell, *The modular difference of classes*, Bulletin of the American Mathematical Society, vol. 23 (1917), pp. 446–450. For a treatment of $x \circ y$ as an analogue of "distance," see M. H. Stone, *Postulates for Boolean algebras and generalized Boolean algebra*, American Journal of Mathematics, vol. 57 (1935), pp. 703–732.

Proof. It is easily seen that when the expression $f(x_1, \cdots, x_n) \circ f(y_1, \cdots, y_n)$ is expanded, all terms where x_i and y_i are primed or not primed together (such as $x_1 \cdots x_n \cdot y_1 \cdots y_n$ or $x_1' \cdots x_{n-1}' \cdot x_n \cdot y_1' \cdots y_{n-1}' \cdot y_n$) have discriminants $=0$. On the other hand, if we expand the expression $(x_1 \circ y_1) + \cdots + (x_n \circ y_n)$, we see that the expansion will be as follows: All terms where x_i and y_i are primed or not primed together have discriminants $=0$, and all other terms have discriminants $=1$. Hence the theorem is proved.

THEOREM 28. *Every Boolean function is uniformly continuous.*

Proof. Let $\epsilon \neq 0$ be arbitrary, and choose $\delta_1 = \cdots = \delta_n = \epsilon$. Then, by Theorem 25, when

$$(x_1 \circ y_1 < \delta_1), \cdots, (x_n \circ y_n < \delta_n),$$

we have

$$f(x_1, \cdots, x_n) \circ f(y_1, \cdots, y_n) < (x_1 \circ y_1) + \cdots + (x_n \circ y_n) < \delta_1 + \cdots + \delta_n = \epsilon + \cdots + \epsilon = \epsilon.$$

11. **Remarks on continuity.** Since we have now seen that all Boolean functions are uniformly continuous, it would be interesting to inquire what non-Boolean functions, if any, are uniformly continuous. Such an investigation, however, would lie outside the bounds I have set myself. I shall merely remark here that one can construct examples to show that some, but not all non-Boolean functions are uniformly continuous. From this it seems likely that the continuity concept might be of importance in the study of a more general type of functions. The relation of Boolean functions to uniformly continuous functions in Boolean algebra is somewhat similar to the relation of polynomials to uniformly continuous functions in ordinary algebra.

12. **One-valued inverses of functions.** When z is a Boolean function of x, then it is possible to consider x as a function of z, say $g(z)$, where $g(z)$, however, is not, in general, one-valued. Similarly, if z is a function of $x_1, \cdots, x_n$, then, again, we may consider x_1 as a function of $z, x_2, \cdots, x_n$ where g, again, may be many-valued.

It is of some interest to investigate the conditions on the given function in order that the inverse function be single-valued. More generally, it is of interest to find, for any f, what domains are such that for x in them f has a single-valued inverse. This question has been solved by Schmidt for a special case: namely, he has shown that a Boolean function of one variable, $f(x) = ax + bx'$, has a one-valued inverse for x in $(ab, a+b)$ and also for x in $(0, a \circ b)$. I propose to find *all* the domains for a function of n variables. I first lay down the following formal definition.

DEFINITION 7. A Boolean function $f(x_1, \cdots, x_n)$ is said to have a *one-valued inverse with respect to x_1 in the domain (A, B)* if, for every u and v in (A, B), we have

$$[f(u, x_2, \cdots, x_n) = f(v, x_2, \cdots, x_n)] \supset (u = v),$$

independent of $x_2, \cdots, x_n$. If $f(x_1, \cdots, x_n)$ has a one-valued inverse for x_1 in the domain $(0, 1)$, we say that $f(x_1, \cdots, x_n)$ has a *one-valued inverse with respect to x_1 everywhere.*

Preliminary to establishing the principal theorem of this section, it is convenient to prove the following lemma.

LEMMA TO THEOREM 29. *If every solution of $a(v_1 \circ v_2)=0$ is also a solution of $b(v_1 \circ v_2)=0$, then $b<a$.*

Proof. Suppose that every solution of

$$a(v_1 \circ v_2) = 0 \tag{1}$$

is also a solution of

$$b(v_1 \circ v_2) = 0. \tag{2}$$

A particular solution of (1) is $v_1=a$, $v_2=1$, since $a(a \circ 1)=aa'=0$. Hence, this must also be a solution of (2); hence $b(a \circ 1)=ba'=0$. Hence $b<a$.

THEOREM 29. *A function $f(x)=ax+bx'$ has a one-valued inverse in those and only those domains that are of the form*

$$(A, A + B[a \circ b]).$$

Proof. First suppose that y_1 and y_2 lie in a domain $(A, A+B[a \circ b])$. Then there exists a v_1 and a v_2 such that

$$y_1 = Av_1' + (A + B[a \circ b])v_1, \quad y_2 = Av_2' + (A + B[a \circ b])v_2. \tag{1}$$

From (1), we see that

$$y_1 \circ y_2 = A'B(a \circ b)(v_1 \circ v_2). \tag{2}$$

Now,

$$f(y_1) \circ f(y_2) = (a \circ b)(y_1 \circ y_2). \tag{3}$$

Thus, from (2) and (3),

$$f(y_1) \circ f(y_2) = y_1 \circ y_2. \tag{4}$$

Hence, if $f(y_1)=f(y_2)$, then $y_1=y_2$; hence $f(x)$ has a one-valued inverse in $(A, A+B[a \circ b])$.

Suppose, on the other hand, that the domain $(A, A+B)$ is such that within it $f(x)$ has a one-valued inverse. Let y_1 and y_2 be any elements in $(A, A+B)$; then there exists a v_1 and a v_2 such that

$$y_1 = Av_1' + (A + B)v_1, \qquad y_2 = Av_2' + (A + B)v_2. \tag{5}$$

By hypothesis,

$$[f(y_1) = f(y_2)] \supset (y_1 = y_2), \tag{6}$$

which is equivalent to

$$[(a \circ b)(y_1 \circ y_2) = 0] \supset [y_1 \circ y_2 = 0]. \tag{7}$$

Substituting (5) in (7) and simplifying, we have

$$[(a \circ b)A'B(v_1 \circ v_2) = 0] \supset [A'B(v_1 \circ v_2) = 0]. \tag{8}$$

Applying the lemma to (8), we have

$$A'B < A'B(a \circ b). \tag{9}$$

But (9) is equivalent to

$$B = AB + B(a \circ b). \tag{10}$$

Thus

$$(A, A + B) = (A, A + AB + B[a \circ b]) = (A, A + B[a \circ b]).$$

It will be noticed that the two results previously mentioned as having been given by Schmidt are special cases of Theorem 29. To show that $f(x)$ has a one-valued inverse in $(ab, a+b)$, set $A=ab$ and $B=1$. To show that $f(x)$ has a one-valued inverse in $(0, a \circ b)$, set $A=0$ and $B=1$. It is evident that the most inclusive domains are got by taking $B=1$.

I now generalize Theorem 29 to functions of n variables.

THEOREM 30. *A function $f(x_1, \cdots, x_n)$ has a one-valued inverse with respect to x_1 in those and only those domains that are of the form*

$$\left(A, A + B \prod_{\alpha_i=0,1} [f(1, \alpha_2, \cdots, \alpha_n) \circ f(0, \alpha_2, \cdots, \alpha_n)]\right).$$

Proof. For $n=1$, we have Theorem 29.

Suppose the theorem true for $n=k$. Let $f(x_1, \cdots, x_k, x_{k+1})$ be any function of $k+1$ variables. Then for any fixed value of x_{k+1}, as y, $f(x_1, \cdots, x_k, y)$ is a function of k variables. Hence by the induction hypothesis we have: $f(x_1, \cdots, x_k, y)$ has a one-valued inverse with respect to x_1 in the domain $(A, A+C)$ if and only if there is a B such that

$$(1) \qquad A + C = A + B \prod_{\alpha_i=0,1} [f(1, \alpha_2, \cdots, \alpha_k, y) \circ f(0, \alpha_2, \cdots, \alpha_k, y)].$$

But there is a B such that (1) holds if and only if

$$(2) \qquad C < A + \prod_{\alpha_i=0,1} [f(1, \alpha_2, \cdots, \alpha_k, y) \circ f(0, \alpha_2, \cdots, \alpha_k, y)].$$

Thus, a necessary and sufficient condition that $f(x_1, \cdots, x_k, y)$ have a one-valued inverse with respect to x_1 in $(A, A+C)$ is that (2) hold. Hence, a necessary and sufficient condition that $f(x_1, \cdots, x_k, x_{k+1})$ have a one-valued inverse with respect to x_1 in $(A, A+C)$ is that (2) hold for all y. But to say that (2) holds for all y is equivalent to saying it holds for $y=1$ and $y=0$, hence equivalent to

$$(3) \qquad \begin{aligned} C &< A + \prod_{\alpha_i=0,1} [f(1, \alpha_2, \cdots, \alpha_k, 1) \circ f(0, \alpha_2, \cdots, \alpha_k, 1)], \\ C &< A + \prod_{\alpha_i=0,1} [f(1, \alpha_2, \cdots, \alpha_k, 0) \circ f(0, \alpha_2, \cdots, \alpha_k, 0)], \end{aligned}$$

which is equivalent to

$$(4) \qquad C < A + \prod_{\alpha_i=0,1} [f(1, \alpha_2, \cdots, \alpha_{k+1}) \circ f(0, \alpha_2, \cdots, \alpha_{k+1})].$$

But to say that (4) holds, is equivalent to saying that there exists a B such that

$$(5) \qquad A + C = A + B \prod_{\alpha_i=0,1} [f(1, \alpha_2, \cdots, \alpha_{k+1}) \circ f(0, \alpha_2, \cdots, \alpha_{k+1})].$$

Hence, the theorem holds for $n=k+1$ if it holds for $n=k$.

THEOREM 31. *A necessary and sufficient condition that a function $f(x_1, \cdots, x_n)$ have a one-valued inverse with respect to x_1 everywhere is that*

$$f(1, \alpha_2, \cdots, \alpha_n) = f'(0, \alpha_2, \cdots, \alpha_n) \qquad [\alpha_i = 0, 1].$$

Proof. Since every domain in which $f(x_1, \cdots, x_n)$ has a one-valued inverse with respect to x_1 is, by Theorem 30, of the form

$$\left(A, A + B \prod_{\alpha_i=0,1} [f(1, \alpha_2, \cdots, \alpha_n) \circ f(0, \alpha_2, \cdots, \alpha_n)]\right),$$

it is clear that if $f(x_1, \cdots, x_n)$ has a one-valued inverse with respect to x_1 in the domain $(0, 1)$, we must take $A=0$ and

$$A + B \prod_{\alpha_i=0,1} [f(1, \alpha_2, \cdots, \alpha_n) \circ f(0, \alpha_2, \cdots, \alpha_n)] = 1.$$

Hence,

$$f(1, \alpha_2, \cdots, \alpha_n) \circ f(0, \alpha_2, \cdots, \alpha_n) = 1 \qquad [\alpha_i = 0, 1].$$

Hence,

$$f(1, \alpha_2, \cdots, \alpha_n) = f'(0, \alpha_2, \cdots, \alpha_n) \qquad [\alpha_i = 0, 1].$$

THEOREM 32. *Those and only those functions have one-valued inverses with respect to x_1 everywhere that are of the form*

$$f(x_1, \cdots, x_n) = [x_1 \Delta f(1, 1, \cdots, 1)]x_2 \cdots x_n + \cdots + [x_1 \Delta f(1, 0, \cdots, 0)]x_2' \cdots x_n'.$$

This theorem is an immediate consequence of Theorem 31.

The next theorem tells us the form which inverse functions have when they exist.

THEOREM 33. *If $f(x_1, \cdots, x_n) = z$ has a one-valued inverse with respect to x_1 everywhere, this inverse function with respect to x_1 is*

$$x_1 = f(z, x_2, \cdots, x_n).$$

Proof. Under the given hypothesis, we have, by Theorem 31,

$$f(1, 1, \cdots, 1)x_1 \cdots x_n + \cdots + f(1, 0, \cdots, 0)x_1 \cdot x_2' \cdots x_n' + f'(1, 1, \cdots, 1)x_1' \cdot x_2 \cdots x_n + \cdots + f'(1, 0, \cdots, 0)x_1' \cdots x_n' = z, \tag{1}$$

which is equivalent to

$$[z \circ f(1, 1, \cdots, 1)]x_1 \cdots x_n + \cdots + [z \circ f(1, 0, \cdots, 0)]x_1 \cdot x_2' \cdots x_n' + [z \circ f'(1, 1, \cdots, 1)]x_1' \cdot x_2 \cdots x_n + \cdots + [z \circ f'(1, 0, \cdots, 0)]x_1' \cdots x_n' = 0, \tag{2}$$

or, again, to

$$\{[z \circ f(1, 1, \cdots, 1)]x_2 \cdots x_n + \cdots + [z \circ f(1, 0, \cdots, 0)]x_2' \cdots x_n'\}\Delta x_1 = 0. \tag{3}$$

Equation (3) has the unique solution

$$x_1 = [z \Delta f(1, 1, \cdots, 1)]x_2 \cdots x_n + \cdots + [z \Delta f(1, 0, \cdots, 0)]x_2' \cdots x_n' = f(z, x_2, \cdots, x_n),$$

as was to be shown.

THE UNIVERSITY OF CALIFORNIA,
BERKELEY, CALIF.

A CONSTRUCTION FOR ABSOLUTE VALUES IN POLYNOMIAL RINGS*

BY

SAUNDERS MacLANE

1. **Introduction.** An absolute value or "Bewertung" of a ring is a function $\|b\|$ which has some of the formal properties of the ordinary absolute value. More explicitly, for any b in the ring, $\|b\|$ must be a real number with the properties†

$$\|bc\| = \|b\| \cdot \|c\|, \qquad \|b + c\| \leqq \|b\| + \|c\|.$$

If the second inequality holds also in the stronger sense

$$\|b + c\| \leqq \max(\|b\|, \|c\|)$$

then the value $\|b\|$ is called non-archimedean (Ostrowski [17], p. 272). The thus delimited non-archimedean values are of considerable arithmetic interest. They are useful in questions of divisibility and irreducibility and in fact often correspond exactly to the prime ideals of the given ring. This paper is devoted to the explicit construction of non-archimedean values. More specifically, given all such values for the field R of rational numbers, we construct all possible values of the ring $R[x]$ of all polynomials in x with coefficients in R.

In treating a non-archimedean value it is convenient to replace $\|a\|$ by a related "exponential" value

$$Va = -\log \|a\|,$$

with corresponding forms (§2) of the formal properties of V (Krull [13], p. 531, and [14], p. 164). All possible non-archimedean values of the field of rational numbers have been determined by Ostrowski ([17], pp. 273–274). For every prime p there is a *p-adic* exponential value V_0 in which the value of any rational number is obtained by writing the number as $p^\alpha(u/v)$, where u and v are prime to p, and setting

$$V_0[p^\alpha(u/v)] = k\alpha, \tag{1}$$

where k is any fixed positive constant. This value we denote by the symbol $[V_0 p = k]$. The only other value V is a trivial one, in which Va is zero for $a \neq 0$.

* Presented to the Society, April 20, 1935, under the title, *Abstract absolute values and polygonal irreducibility criteria*; received by the editors June 28, 1935.

† See Kürschák [15]. (Numbers in brackets refer to the bibliography at the end of the paper.)

On this basis we can determine all possible values in the ring of polynomials with rational coefficients. Any such value W gives a p-adic or trivial value $V_0 a = Wa$ for the rational numbers and a value $\mu = Wx$ for the variable x. These facts alone give a first approximation V_1 to the value W, as follows:

$$V_1(a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0) = \min\,[V_0 a_n + n\mu,\ V_0 a_{n-1} + (n-1)\mu,\ \cdots,\ V_0 a_0]. \tag{2}$$

This V_1 is actually a value and is never larger than W. If V_1 is not equal to W, we choose a $\phi(x)$ of smallest possible degree for which $W\phi(x) > V_1\phi(x)$. We then define a second approximation $V_2 f(x)$ by using* the true value for $\phi(x)$. In this manner we construct successive approximations $V_1, V_2, V_3, \cdots$ which in the limit will give the arbitrary value W (§8).

The succession of values $V_1, V_2, \cdots$ is defined in Part I for polynomials with coefficients in any field K. This requires a general method (§§4 and 5) of constructing a value V_k from a previously obtained value V_{k-1}. The value given by the limit of such a sequence needs a special study (§7). Here, as in §§8 and 16, we assume that every value of the field K is "discrete"; that is, that the real numbers used as values form an isolated point set, as in the case of p-adic values.

Part II investigates the structure of the values which have been constructed. The central problem is the construction of the "residue-class field" which arises when polynomials which differ by a polynomial of positive value are put into the same residue-class. For the absolute values constructed in Part I this field is determined by an inductive construction of the homomorphism of polynomials to residue-classes (§§10–14). This homomorphism also yields a more specific description of how our values can be built up (§§9, 13). Since a given value W can be represented in many ways by a sequence of approximations $V_1, V_2, V_3, \cdots$, we treat in §§15 and 16 the questions as to when two such sequences can give the same ultimate value W, and how such a sequence can be put in a normal form.

Among the applications of this construction of absolute values we mention the classification of irreducibility criteria of the Newton Polygon type. The theorem of Eisenstein [4] states that a polynomial

$$f(x) = x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_0$$

with integral coefficients a_i is irreducible if each coefficient a_i is divisible by some fixed prime p, while the last term a_0 is not divisible by p^2. In terms of the value V_1 of (2) with $\mu = 1/n$ these hypotheses on $f(x)$ become

* Similar "second-stage" values V_2 appear implicitly in the irreducibility investigations of Ore [7], Kürschák [6], and Rella [10].

$$V_1 f(x) = V_1 x^n = V_1 a_0 < V_1(a_i x^i) \qquad (i = 1, \cdots, n-1).$$

In this form a simple proof of the theorem can be given. The theorems of Königsberger [5], Dumas [3], and Ore [8] are likewise related to the values V_1. The second stage values V_2 can be similarly applied to interpret the irreducibility theorems of Schöneman [11], Bauer [1], Kürschák [6], and Ore [7]. By using the general value V_k one can obtain a still more extensive irreducibility criterion which includes all these previous theorems (MacLane [16]), and which asserts that certain polynomials $f(x)$ with irreducible homomorphic images of sufficiently high degree are themselves irreducible. Our construction for absolute values can also be applied to give a new and complete treatment of the problem of constructing the prime ideal factors of a given rational prime in a given algebraic field.*

I. The construction of non-archimedean values

2. **Elementary properties of values in rings.** A ring† S is said to have a *non-archimedean value* (for short, a value) V if to every element $a \neq 0$ in S there is assigned a unique real number Va with the properties

$$V(ab) = Va + Vb, \qquad V(a+b) \geqq \min(Va, Vb).$$

These we call the *product* and *triangle* laws respectively. We assume also that 0 is assigned the value $+\infty$, with the following conventions for any finite number γ:

$$\gamma < \infty, \qquad \infty + \gamma = \gamma + \infty = \infty + \infty = \infty.$$

Two simple consequences of the product law are

$$V(1) = V(-1) = 0, \qquad V(-a) = V(a). \tag{1}$$

More important is the strengthened form of the triangle law:

$$Va \neq Vb \text{ implies } V(a+b) = \min(Va, Vb). \tag{2}$$

For suppose instead that $Va > Vb$ and $V(a+b) > \min(Va, Vb)$. Then

$$Vb = V(a+b-a) \geqq \min(V(a+b), Va) > Vb,$$

a contradiction.

Since we are using a value analogous to the negative logarithm of the ordinary absolute value, a "small" absolute value will correspond to a "large"

* The application of the methods of this paper to the prime ideal construction is treated in another paper (Duke Mathematical Journal, vol. 3 (1936), pp. 492–510.) The general irreducibility criterion will be presented in a subsequent paper.

† Here and in the sequel "ring" means "commutative ring with unit element."

value V. Hence we say that two ring elements a and b are of the same order of magnitude or *equivalent* in* V—symbol $a \infty b$(in V)—if and only if

$$V(a - b) > Va.$$

The product and strong triangle laws show that equivalent elements have the same value and that equivalence is a reflexive, symmetric, and transitive relation, provided the supplementary assumption† that $0 \infty 0$ be made. Two equivalences

$$a \infty b \quad \text{and} \quad c \infty d$$

can be multiplied to give

$$ac \infty bd. \tag{3}$$

An element b is *equivalence-divisible* in V by a if and only if there exists a c in S such that

$$b \infty ca \quad (\text{in } V).$$

If this is true, it remains true when a or b is replaced by an equivalent element.

The product law implies that a ring S with a value V must be a domain of integrity. The value V may be extended to the quotient field of S by defining, in accord with the product law,

$$V\left(\frac{a}{b}\right) = Va - Vb \tag{4}$$

for any elements a and $b \neq 0$ in S. One then obtains the

THEOREM 2.1. *Let S be a domain of integrity with the quotient field K. If V is a value of S, then the function defined by* (4) *is a value of K. Conversely, every value of K can be obtained in this way from one and only one value of S.*

When $S = K$ is a field, the set of all real numbers Va for $a \neq 0$ in S is an additive group Γ, called the *value-group* of V. If the positive numbers of Γ have a positive minimum $\delta > 0$, then the value V is said to be *discrete*.‡ In this case the group Γ is cyclic and consists of all multiples of δ. If all elements not 0 have the value 0, then V is called *trivial*. Every ring has a trivial value, while the p-adic values for the field of rational numbers are examples of discrete values. Values of arithmetic interest are generally discrete.

* This term, used by Rella [10], is similar to Ore's "congruent modulo a polygon" ([8], p. 270) and Kürschák's "equipollence" ([6], p. 185).

† Here and subsequently the element 0 plays an exceptional role.

‡ Krull [14], p. 171, and Hasse-Schmidt [12], p. 31.

3. **The first stage values.** Our problem is this: Given all values of a field K; to construct all values for the ring $K[x]$ of all polynomials* in x with coefficients in K. By Theorem 2.1, this is equivalent to determining all values in the field $K(x)$ of rational functions of x with coefficients in K. No gain in generality would result were a ring S used instead of the field K.

As indicated in the introduction, the values for $K[x]$ will be constructed in stages. For the first step, take any value V_0 for the field K and any real number μ, and then define a corresponding *first stage value* V_1 for any polynomial by the equation (2) of §1. In particular, this gives

$$V_1 x = \mu, \qquad V_1 a = V_0 a \qquad \text{(any } a \in K\text{)}.$$

Hence we use the symbol $[V_0, V_1 x = \mu]$ for the value V_1.

THEOREM 3.1. *If V_0 is a value of K and μ a real number, the function $V_1 = [V_0, V_1 x = \mu]$ defined above is a value of $K[x]$.*

This has been proved by Rella ([10], pp. 35–36) and by Ostrowski ([18], p. 363). The latter calls x an "invariant element." A particularly simple V_1 arises when $\mu = 0$. On the other hand, if V_0 is trivial and $\mu < 0$, then

$$V_1 a(x) = \mu \cdot \deg a(x).$$

The symbol deg $a(x)$ here and in the sequel denotes the *degree in x* of the polynomial $a(x)$.

4. **Augmented values.** Our construction now proceeds to build a second stage value on the basis of a first stage one; or, more generally, a kth stage value from one at the stage $k-1$. The process involved can be formulated once for all: Given a value W for $K[x]$; to construct an "augmented" value V by assigning larger values to a certain "key" polynomial $\phi(x)$ and to its equivalence-multiples. The *key polynomial* $\phi(x)$ must be suitably chosen.

DEFINITION 4.1. *A key polynomial $\phi(x) \neq 0$ over a value W of $K[x]$ is one which satisfies the following conditions*:

(i) *Irreducibility. If a product is equivalence-divisible in W by $\phi(x)$, then one of the factors is equivalence-divisible by $\phi(x)$.*

(ii) *Minimal degree. Any non-zero polynomial equivalence-divisible in W by $\phi(x)$ has a degree in x not less than the degree of $\phi(x)$.*

(iii) *The leading coefficient† of $\phi(x)$ is* 1.

This key polynomial is to be assigned a new value

$$V\phi(x) = \mu > W\phi(x). \tag{1}$$

* Henceforth all polynomials considered are to have coefficients in K, unless otherwise noted.

† This assumption, although unnecessary, will simplify the subsequent work.

To find the new values of other polynomials, we use *expansions in* ϕ; that is, expressions in powers of $\phi(x)$ of the form*

$$f(x) = f_m(x)\phi^m + f_{m-1}(x)\phi^{m-1} + \cdots + f_0(x), \tag{2}$$

in which each coefficient polynomial $f_i(x)$ is either zero or of degree less than the degree of $\phi(x)$. Any polynomial has one and only one such expansion, which may be found by successive division by powers of ϕ. The new value $Vf(x)$ is computed from the expansion thus:

$$V[f_m(x)\phi^m + f_{m-1}(x)\phi^{m-1} + \cdots + f_0(x)] = \min_i [Wf_i(x) + i\mu]. \tag{3}$$

Here "min" with subscript i means the smallest quantity of the form $Wf_i(x)+i\mu$, for $i=0, 1, \cdots, m$.

THEOREM 4.2. *If W is a value of $K[x]$, $\phi(x)$ is a key polynomial over W and μ is a real number satisfying* (1), *then the function V defined in* (3) *is also a value of $K[x]$. V is called an augmented value, and is denoted by*

$$V = [W, V\phi = \mu].$$

Proof. The product and triangle laws for V must be verified.† We first prove the triangle law for a sum $f(x)+g(x)$. Let f and g have the expansions (2) and

$$g(x) = g_n(x)\phi^n + g_{n-1}(x)\phi^{n-1} + \cdots + g_0(x) \tag{4}$$

respectively. By adjoining zero coefficients we can make $m=n$. Hence $f+g$ has the expansion

$$f(x) + g(x) = \sum_{i=0}^{n} [f_i(x) + g_i(x)]\phi^i.$$

By the definition of V and the triangle law for W,

$$\begin{aligned} V(f+g) &= \min_i [W(f_i + g_i) + i\mu] \geqq \min_i [\min (Wf_i, Wg_i) + i\mu] \\ &\geqq \min_i [Wf_i + i\mu, Wg_i + i\mu] \\ &= \min [\min_i (Wf_i + i\mu), \min_i (Wg_i + i\mu)], \\ V(f+g) &\geqq \min [Vf, Vg]. \end{aligned}$$

---

* We use ϕ as an abbreviation for $\phi(x)$, and similarly for other polynomials.

† The proof resembles one of Rella's ([10], pp. 36–37); the product-law proof is an extension of the Dumas-Kürschák-Ore proof for the Newton polygon of a product.

To prove the product law we will use the *quotient-remainder* expression for a polynomial $f(x)$,

$$f(x) = q(x)\phi + r(x), \tag{5}$$

where $r(x)$ is zero or of degree less than that of $\phi(x)$.

LEMMA 4.3. *If ϕ is a key polynomial over a value W of $K[x]$, and if $f(x) \neq 0$ has the quotient-remainder expression* (5), *then*

$$Wr(x) \geqq Wf(x), \tag{6}$$

$$W(q(x)\phi) \geqq Wf(x). \tag{7}$$

The inequality in (6) *holds if and only if $f(x)$ is equivalence-divisible by ϕ in W.*

Were the first conclusion (6) false, then $Wf(x) > Wr(x)$ in (5) and the definition of equivalence would give

$$r(x) \backsim (-q(x))\phi \qquad (\text{in } W).$$

Hence $r(x) \neq 0$ is equivalence-divisible by ϕ, a contradiction to the minimal property of ϕ and the restricted degree of $r(x)$. The second conclusion (7) now follows from (6) by the triangle law.

The third conclusion gives a test for equivalence-divisibility in terms of ordinary division. When $Wr(x) > Wf(x)$, then (5) shows $f(x)$ equivalence-divisible by ϕ. Conversely, if $f(x)$ is equivalence-divisible in W by ϕ, then there exist polynomials $h(x)$ and $s(x)$ so that

$$f(x) = h(x)\phi + s(x), \qquad Ws(x) > Wf(x).$$

If now the equality sign in (6) should hold, we would have

$$r(x) = f(x) - q(x)\phi = (h(x) - q(x))\phi + s(x),$$

with

$$Ws(x) > Wf(x) = Wr(x),$$

making ϕ an equivalence-divisor of $r(x)$, again a contradiction. The lemma is proved.

Return to Theorem 4.2 and consider the product law first for a product of two monomial expansions $a(x)\phi^t$ and $b(x)\phi^u$. Because of the limited degrees of $a(x)$ and $b(x)$, the product $a(x)b(x)$ has an expansion with not more than two terms,

$$a(x)b(x) = c(x)\phi + d(x). \tag{8}$$

The product $a(x)b(x)$ is not equivalence-divisible in W by ϕ, for if it were, the equivalence-irreducibility of ϕ (Definition 4.1) would require that one of the

factors be equivalence-divisible by ϕ, contrary to the minimal property. Lemma 4.3 and the triangle axiom yield then

$$W(c(x)\phi) \geqq W(a(x)b(x)) = Wd(x).$$

Since the new value of ϕ exceeds the old value,

$$Wc(x) + \mu > W(a(x)b(x)) = Wd(x). \tag{9}$$

The product under consideration has by (8) the expansion

$$(a(x)\phi^t)(b(x)\phi^u) = c(x)\phi^{t+u+1} + d(x)\phi^{t+u};$$

hence the definition of V and the conclusion (9) give

$$\begin{aligned} V[(a(x)\phi^t)(b(x)\phi^u)] &= \min\,[Wc(x) + \mu + (t+u)\mu,\ Wd(x) + (t+u)\mu] \\ &= Wd(x) + (t+u)\mu \\ &= Wa(x) + t\mu + Wb(x) + u\mu \\ &= V(a(x)\phi^t) + V(b(x)\phi^u). \end{aligned}$$

This is the product law for monomial expansions.

The product law for polynomials $f(x)$ and $g(x)$ with arbitrary expansions (2) and (4) respectively is an immediate consequence. The product $f(x)g(x)$ has an expansion obtained by adding expansions of monomial products; hence

$$V(f(x)g(x)) \geqq Vf(x) + Vg(x). \tag{10}$$

To show that the equality holds, choose t and u as the largest integers with

$$V(f_t(x)\phi^t) = Vf(x), \qquad V(g_u(x)\phi^u) = Vg(x)$$

respectively. The monomial case then shows* that the expansion of $f(x)g(x)$ has a term $r(x)\phi^{t+u}$ with the value $Vf+Vg$. The equality holds in (10), and Theorem 4.2 is established.

5. **Properties of augmented values.** An augmented value V is never less than the original value W. This characteristic property will now be established. As a consequence the method used to compute V can be extended (Theorem 5.2) in a way subsequently useful in §12.

THEOREM 5.1. (*Monotonicity.*) *The augmented value* $V = [W, V\phi = \mu]$ *makes*

$$Vf(x) \geqq Wf(x)$$

for all polynomials $f(x) \neq 0$. *The inequality sign holds if and only if* $f(x)$ *is equivalence-divisible in* W *by* ϕ. *In particular, the equality sign holds whenever the degree of* $\phi(x)$ *exceeds that of* $f(x)$.

* The details here omitted are given in Rella's proof.

The proof is by induction on the degree m of the expansion of $f(x)$ in ϕ (see §4, (2)). If $m=0$, the definition of V shows $Vf(x)$ and $Wf(x)$ equal. If $m>0$, the quotient-remainder expression

$$f(x) = q(x)\phi + r(x) \tag{1}$$

indicates that $q(x)$ has an expansion of degree $m-1$ in ϕ; hence the induction assumption will be

$$Vq(x) \geqq Wq(x).$$

The value of the first term on the right of (1), by §4, (1), and the quotient-remainder Lemma 4.3, is

$$V(q(x)\phi) \geqq Wq(x) + V\phi > W(q(x)\phi) \geqq Wf(x).$$

For the second term, the case $m=0$ and Lemma 4.3 imply

$$Vr(x) = Wr(x) \geqq Wf(x),$$

where the inequality holds if and only if $f(x)$ is equivalence-divisible by ϕ in W. The strong triangle law for V applied to (1) now gives the result (see §2, (2)).

THEOREM 5.2. *If in the expression*

$$a(x) = a_n(x)\phi^n + a_{n-1}(x)\phi^{n-1} + \cdots + a_0(x)$$

the degrees of the $a_i(x)$ are not limited, but if no $a_i(x)$ is equivalence-divisible in W by ϕ, then the augmented value $V=[W, V\phi=\mu]$ is

$$Va(x) = \min_i [Wa_i + i\mu] \qquad (i = 0, 1, \cdots, n).$$

Proof. A quotient-remainder expression for each coefficient polynomial gives

$$a_i(x) = q_i(x)\phi + r_i(x) \qquad (i = 0, 1, \cdots, n),$$

$$a(x) = \sum_{i=0}^{n} q_i(x)\phi^{i+1} + \sum_{i=0}^{n} r_i(x)\phi^i.$$

Lemma 4.3 shows that the second summation has a value

$$V\Big(\sum_i r_i\phi^i\Big) = \min_i [Wr_i + i\mu] = \min_i [Wa_i + i\mu]$$

and that the first summation has a larger value

$$V\Big(\sum_i q_i\phi^{i+1}\Big) \geqq \min_i [Vq_i + V\phi + i\mu] > \min_i [Wq_i + W\phi + i\mu],$$

because of the monotonicity. The strong triangle law for the sum of these two summations yields the desired conclusion.

6. **Inductive and limit-values.** This section classifies the values and value-groups obtained by successive augmented values.

DEFINITION 6.1. *A kth stage inductive value V_k is any value of $K[x]$ obtained by a sequence of values $V_1, V_2, \cdots, V_k$, where $V_1 = [V_0, V_1 x = \mu_1]$ is a first stage value* (§3) *and where each V_i is obtained by augmenting V_{i-1};*

$$V_i = [V_{i-1}, V_i\phi_i = \mu_i] \qquad (i = 2, 3, \cdots, k).$$

*Furthermore, for $i = 2, \cdots, k$, the key polynomials $\phi_i(x)$ must satisfy:**

(1) $\deg \phi_i(x) \geqq \deg \phi_{i-1}(x)$;

(2) *the equivalence $\phi_i(x) \backsim \phi_{i-1}(x)$ (in V_{i-1}) is false.*

Here the first key polynomial is understood to be $\phi_1(x) = x$.

This value V_k may be conveniently symbolized thus:

$$V_k = [V_0, V_1 x = \mu_1, V_2\phi_2 = \mu_2, V_3\phi_3 = \mu_3, \cdots, V_k\phi_k = \mu_k]. \tag{3}$$

Given an infinite sequence $V_1, V_2, \cdots, V_k, \cdots$ of such values, we set

$$V_\infty f(x) = \lim_{k \to \infty} V_k f(x). \tag{4}$$

The monotonic character of V_k indicates that this limit, if not finite, is $+\infty$. V_∞ satisfies the product law for values, as can be shown by taking limits in the product law for V_k. As for the sum $f(x)+g(x)$, note that the triangle law in V_k indicates that one of the inequalities

$$V_k(f(x) + g(x)) \geqq V_k f(x), \qquad V_k(f(x) + g(x)) \geqq V_k g(x)$$

holds for infinitely many k. One of the conclusions

$$V_\infty(f(x) + g(x)) \geqq V_\infty f(x), \qquad V_\infty(f(x) + g(x)) \geqq V_\infty g(x)$$

then results, and thence follows the triangle law for V_∞. We have

THEOREM 6.2. *Let $\{\phi_k(x)\}$ and $\{\mu_k\}$ be fixed infinite sequences such that all the functions V_k indicated in* (3) *are inductive values. Then the function $V_\infty f(x)$ defined in* (4) *is a value of $K[x]$, provided some polynomials not zero be allowed to have the value $+\infty$.*

This function V_∞ will be called a *limit-value*. The case when several successive key polynomials have the same degree will often require separate treatment, based on

* These conditions involve no loss of generality, but simplify subsequent proofs (see Theorem 6.7 and the end of §9).

LEMMA 6.3. *If in the inductive value V_k in* (3) *the key polynomials $\phi_{t+1}(x)$, $\phi_{t+2}(x), \cdots, \phi_k(x)$ all have the same degree, for t with $0 \leqq t < k-1$, then*

(i) $V_t(\phi_{j+1} - \phi_j) = \mu_j$ $\quad (j = t+1, t+2, \cdots, k-1)$,

(ii) $\mu_k > \mu_{k-1} > \cdots > \mu_{t+1}$,

(iii) $V_t\phi_k = V_t\phi_{k-1} = \cdots = V_t\phi_{t+1}$ $\quad$ *(if $t>0$).*

Proof. Let j range from $t+1$ to $k-1$, and set

$$s_j(x) = \phi_{j+1}(x) - \phi_j(x). \tag{5}$$

Since both ϕ's have the first coefficient 1, the degree of $s_j(x)$ is less than that of $\phi_j(x)$. Therefore, by Theorem 5.1,

$$V_t s_j(x) = V_{t+1} s_j(x) = \cdots = V_k s_j(x).$$

If the first conclusion were false for some j, we would have

$$V_t s_j(x) = V_j(\phi_{j+1} - \phi_j) > \mu_j = V_j\phi_j,$$

for the other inequality is impossible by Lemma 4.3. This would give

$$\phi_{j+1} \backsim \phi_j \quad (\text{in } V_j),$$

a contradiction of assumption (2). The conclusion (i) is thus established. Coupled with the monotonicity and the triangle axiom for (5), it gives the second conclusion, for

$$\mu_{j+1} = V_{j+1}\phi_{j+1} > V_j\phi_{j+1} \geqq \min [V_j\phi_j, V_t s_j] = \mu_j.$$

For similar reasons, assuming now that $t>0$,

$$V_t s_j(x) = \mu_j = V_j\phi_j > V_{j-1}\phi_j \geqq V_t\phi_j.$$

The strong triangle axiom for V_t in (5) then yields conclusion (iii),

$$V_t\phi_{j+1} = \min [V_t s_j(x), V_t\phi_j] = V_t\phi_j.$$

An interesting consequence of this lemma is the invariance of the values assigned to the key polynomials.

THEOREM 6.4. *If the ith stage of the inductive value V_k in* (3) *uses a key polynomial ϕ_i with an assigned value μ_i, then*

$$V_k\phi_i(x) = V_i\phi_i(x) = \mu_i.$$

For this conclusion follows directly from Theorem 5.1 if the degree of $\phi_{i+1}(x)$, and hence that of every subsequent key polynomial, exceeds the degree of $\phi_i(x)$. The only case remaining is that of Lemma 6.3, with $t=i-1$. But, by (5),

$$\phi_i = \phi_k - s_{k-1}(x) - s_{k-2}(x) - \cdots - s_i(x).$$

The terms on the right have by the preceding lemma the V_k values $\mu_k, \mu_{k-1}, \cdots, \mu_i$ respectively, so that the conclusion follows by the strong triangle law. Both this theorem and Lemma 6.3 hold equally well for limit-values.

The monotonic property of inductive values can be sharpened thus:

THEOREM 6.5. *Let a limit or inductive value be built up by the inductive values $V_1, V_2, \cdots$. Then, for any fixed polynomial $f(x) \neq 0$, either*

$$V_{k+1}f(x) > V_kf(x) \qquad (k = 1, 2, \cdots),$$

or else there is an $i \geqq 1$ such that

$$V_1f(x) < V_2f(x) < \cdots < V_{i-1}f(x) < V_if(x) = V_{i+1}f(x) = V_{i+2}f(x) = \cdots.$$

In the latter case there is an $r(x)$ of degree less than that of ϕ_{i+1} with

$$f(x) \infty r(x) \qquad (in\ V_k) \quad (k = i+1, i+2, \cdots).$$

Suppose, contrary to the first alternative, that for some i

$$V_{i+1}f(x) = V_if(x).$$

Then the quotient-remainder expression

$$f(x) = q(x)\phi_{i+1} + r(x)$$

must by Theorem 5.1 and Lemma 4.3 have $V_ir = V_if$. Hence, for any $k \geqq i+1$,

$$V_k(f - r) \geqq V_{i+1}(f - r) = V_{i+1}(q\phi_{i+1}) > V_i(q\phi_{i+1}) \geqq V_if = V_ir = V_kr.$$

Therefore $f(x)$ and $r(x)$ are equivalent in V_k and

$$V_kf = V_kr = V_ir = V_if,$$

so that $V_kf(x)$ is constant for $k \geqq i$, which is the second alternative.

An inductive value V_k of $K[x]$ gives by Theorem 2.1 a value for the field $K(x)$ of rational functions. This value has by §2 a value-group Γ_k, which we call the *value-group associated with* V_k. It may be determined in the following way:

THEOREM 6.6. *The value V_k in* (3) *has a value-group Γ_k consisting of all real numbers of the form*

$$v + m_1\mu_1 + m_2\mu_2 + \cdots + m_k\mu_k,$$

where the m_i are integers and v is an element of the value-group of the original value V_0.

That every number of Γ_k must be of this form follows by induction from the definition of the augmented value V_k. Conversely, any number of this form is by Theorem 6.4 the value in V_k of the rational function

$$bx^{m_1}\phi_2^{m_2}\cdots\phi_k^{m_k},$$

where b is a constant in K with the value ν.

For a more precise description, designate a real number μ as *commensurable* with an additive group of numbers whenever some integral multiple of μ lies in the group (Ostrowski [18], p. 367). Then

THEOREM 6.7. *In an inductive value V_k from* (3) *every assigned value μ_i, except perhaps μ_k, is commensurable with the value-group Γ_{i-1} of the preceding value* (*the case $i=1$ included*).

For a proof, consider the expansion in ϕ_i of the next key,

$$\phi_{i+1}(x) = f_m(x)\phi_i^m + f_{m-1}(x)\phi_i^{m-1} + \cdots + f_0(x).$$

If μ_i is not commensurable with Γ_{i-1}, no two terms here can have the same value in V_i. Only one term, say the jth, has the minimum value, and

$$\phi_{i+1}(x) \backsim f_j(x)\phi_i^j \qquad (\text{in } V_i).$$

By the irreducibility of ϕ_{i+1} at least one of the conditions:

(6) $f_j(x)$ is equivalence-divisible in V_i by ϕ_{i+1},

(7) $\phi_i(x)$ is equivalence-divisible in V_i by ϕ_{i+1},

must hold. Because of the minimal property of ϕ_{i+1} the first possibility (6) contradicts the assumption (1) of Definition 6.1. For the same reasons the second possibility (7) implies that ϕ_{i+1} and ϕ_i have the same degree, while

$$s(x) = \phi_{i+1}(x) - \phi_i(x)$$

has a smaller degree. Because of (7), Lemma 4.3 applied to V_i and the key polynomial ϕ_{i+1} shows $V_i s(x) > V_i\phi_i(x)$. Hence

$$\phi_i(x) \backsim \phi_{i+1}(x) \qquad (\text{in } V_i),$$

a contradiction of assumption (2). There can be no next key ϕ_{i+1}.

7. **Constant degree limit-values.** A limit-value V_∞ for polynomials does not give a value for all rational functions if some of the polynomials have the value $+\infty$. Hence the problem: When is V_∞ *finite*; that is, when is $V_\infty f(x)$ finite for all $f(x)\neq 0$? We obtain an answer in the discrete case.

If the key polynomials $\phi_k(x)$ increase indefinitely in degree, then $V_k f(x)$ is by Theorem 5.1 ultimately constant for fixed $f(x)$ and V_∞ is finite. A different situation arises if the degrees of $\phi_k(x)$ have an upper bound. By assumption (1) of Definition 6.1, the degrees of $\phi_k(x)$ are then all equal to some M for k sufficiently large. For an example of such a *constant degree limit-value*, start with the p-adic value $[V_0 3 = 1]$ for the rational field (see §1, (1)) and set

$$V_1 = [V_0, V_1 x = 1],$$
$$V_k = [V_{k-1}, V_k(x + 2p + p^2 + p^3 + \cdots + p^{k-1}) = k] \qquad (k = 2, 3, \cdots).$$

This gives a limit-value of constant degree 1. Since

$$\frac{p}{2} = 2p + p^2 + p^3 + \cdots + p^{k-1} - \frac{p^k}{2} \qquad (k > 1, p = 3)$$

holds by the usual methods for p-adic numbers, we find

$$V_\infty\left(x + \frac{p}{2}\right) = \lim_{k\to\infty} V_k\left[(x + 2p + p^2 + \cdots + p^{k-1}) - \frac{p^k}{2}\right] = \lim_{k\to\infty} k = \infty.$$

Hence this V_∞ is not finite.

This use of p-adic numbers suggests the general notion of a perfect ring. In any ring S with a value V, a sequence $\{a_n\}$ is a *Cauchy sequence* if $V(a_n - a_m)$ approaches ∞ with n and m. If every Cauchy sequence has a V-limit b such that $V(a_n - b)$ approaches ∞ with n, the ring S is said to be *perfect*. Any ring can be embedded in a perfect ring by the usual procedure of adjoining limits of Cauchy sequences (Kürschák [15] and Hasse-Schmidt [12], p. 24).

THEOREM 7.1. (*Finiteness criterion.*) *Let V_∞ be a limit-value with key polynomials $\phi_k(x)$ of constant degree M for $k > t > 0$. Extend the ring $K[x]$ with the value V_t to be a perfect ring S^*. Assume that all values of K are discrete. Then $\{\phi_k\}$ is a Cauchy sequence in V_t and has a limit ϕ in S^*. Furthermore V_∞ is finite if and only if there is no $g(x) \neq 0$ in $K[x]$ divisible in S^* by the limit ϕ.*

For V_∞ the symbolism of Theorem 6.2 may be used. Since $\phi_{t+1}, \phi_{t+2}, \cdots$ all have the same degree M, the conclusions of Lemma 6.3 on constant degree values are applicable. Each number μ_i is by Theorem 6.7 commensurable with the value-group Γ_{i-1} of V_{i-1}. Our assumption shows the original value-group Γ_0 of V_0 to be discrete, hence, by Theorem 6.6 and by induction, the group Γ_t is discrete. But Lemma 6.3 gives

$$\mu_i = V_t(\phi_{i+1} - \phi_i) \in \Gamma_t \qquad (i > t); \tag{1}$$

hence $\Gamma_i = \Gamma_t$ for $i > t$. This lemma also shows the sequence $\{\mu_i\}$ to be monotone increasing for $i > t$; it lies in the discrete set Γ_t, hence

$$\lim_{i\to\infty} \mu_i = \infty. \tag{2}$$

The strong triangle law combined with (1) then proves

$$V_t(\phi_{i+j} - \phi_i) = V_t\left[\sum_{k=i}^{i+j-1} (\phi_{k+1} - \phi_k)\right] = \min_k [\mu_k] = \mu_i.$$

Therefore, by (2), $\{\phi_i\}$ is a Cauchy sequence with a limit ϕ in S^*. This ϕ need not be a polynomial, but, by conclusion (iii) of Lemma 6.3, $\phi \neq 0$.

Now consider the necessary condition for finiteness. If $g(x) \neq 0$ is divisible by ϕ in S^*, then

$$g(x) = h \cdot \phi,$$

where h is the V_t-limit of a Cauchy sequence $\{h_i(x)\}$ from $K[x]$. The usual argument for the convergence of a product shows

$$\lim_{i \to \infty} V_t[g(x) - h_i(x) \cdot \phi_i(x)] = \infty. \tag{3}$$

By the triangle axiom and the monotonic property for $i > t$,

$$V_i g \geqq \min [V_i(h_i\phi_i), V_i(g - h_i\phi_i)] \geqq \min [V_t h_i + \mu_i, V_t(g - h_i\phi_i)]. \tag{4}$$

But $\{h_i(x)\}$ is a convergent sequence in V_t with a limit not zero, so that, as is well known ([12], p. 25), $V_t h_i$ is ultimately constant. Consequently (2), (3), and (4) prove

$$V_\infty g(x) = \lim_{i \to \infty} V_i g(x) = \infty, \tag{5}$$

so that V_∞ is not a finite limit-value.

Conversely, suppose that V_∞ is not finite. Then (5) holds for some $g(x) \neq 0$. If $g(x)$ has the quotient-remainder expressions $q_i(x)\phi_i + r_i(x)$, then, by Theorem 5.1 and by Lemma 4.3,

$$V_t(g - q_i\phi_i) = V_t r_i = V_{i-1} r_i \geqq V_{i-1} g(x) \to \infty \qquad (i > t).$$

Thus the sequence $\{q_i\phi_i\}$ converges in V_t to the limit $g(x) \neq 0$. Since $\{\phi_i\}$ already converges to the limit $\phi \neq 0$, the standard argument for the limit of a quotient $(q_i\phi_i)/\phi_i$ shows that $\{q_i\}$ must converge in V_t to some limit q in S^*, such that

$$f(x) = q\phi.$$

Hence ϕ is a factor of $f(x)$ in S^*, as asserted.

8. **Completeness.** We have the following theorem.

THEOREM 8.1. *If every value of the field K is discrete, then every non-archimedean value W of the ring $K[x]$ can be represented either as an inductive or as a limit-value.*

Given W, we shall construct by stages a corresponding inductive value V_k with the following three properties (notation as in §6, (3)):

(1) $$Wf(x) \geqq V_k f(x) \qquad \text{(for all } f(x)\text{)},$$

(2) $$\deg f(x) < \deg \phi_k \text{ implies } Wf(x) = V_k f(x),$$

(3) $$W\phi_i(x) = V_k\phi_i(x) = \mu_i \qquad (i = 1, 2, \cdots, k).$$

The initial value V_1 is defined by

$$\mu_1 = Wx, \qquad V_0 a = Wa \qquad \text{(any } a \in K\text{)};$$

the triangle axiom for W and the definition of V_1 in §1, (2), then show that conditions (1), (2), and (3) hold for $k=1$.

Suppose now that an inductive value V_k with these three properties has already been constructed, and that the equality in (1) does not always hold. As a prospective key polynomial, choose a $\psi(x)$ of smallest possible degree with the property

(4) $$W\psi(x) > V_k\psi(x).$$

Multiplication with some constant gives $\psi(x)$ the first coefficient 1. Furthermore the two statements,

(5) $$Wf(x) > V_k f(x),$$

(6) $$f(x) \text{ is equivalence-divisible in } V_k \text{ by } \psi(x),$$

are logically equivalent. For if (5) is given, and if $f(x)$ has the quotient-remainder expression $q(x)\psi+r(x)$, then

$$V_k(q\psi - f) = W(q\psi - f) \geqq \min\,[W(q\psi), Wf] > \min\,[V_k(q\psi), V_k f],$$

because of (2), the minimum degree choice of ψ and the induction assumption (1) for $q(x)$. Hence the strong triangle law shows $f \infty q\psi$ in V_k, which is the conclusion (6). Conversely, if (6) holds there exist polynomials $h(x)$ and $s(x)$ with

$$f(x) = h(x)\psi + s(x), \qquad V_k s(x) > V_k f(x) = V_k(h(x)\psi).$$

Then, because of the induction assumption (1),

$$Wf \geqq \min\,[W(h\psi), Ws] \geqq \min\,[V_k h + W\psi, V_k s] > V_k h + V_k\psi = V_k f,$$

which gives conclusion (5). The equivalence of (5) and (6) is established.

From the equivalence one readily shows that $\psi(x)$ satisfies the Definition 4.1 of a key polynomial over the value V_k. Finally we can assign $\psi(x)=\phi_{k+1}$ the new value

(7) $$\mu_{k+1} = W\psi > V_k\psi,$$

satisfying the proper inequality, and then construct the augmented value $V_{k+1}= [V_k, V_{k+1}\phi_{k+1}=\mu_{k+1}]$. This will be an inductive value if only conditions

(1) and (2) of Definition 6.1 hold. By the choice of $\phi_{k+1}=\psi$ and the induction assumption (2), $\phi_k(x)$ cannot exceed $\phi_{k+1}(x)$ in degree, therefore condition (1) of §6 is true. Condition (2) of §6 could only be false if ϕ_{k+1} and ϕ_k were equivalent in V_k; in other words, only if

$$V_k(\phi_k - \phi_{k+1}) > V_k\phi_k = V_k\phi_{k+1}.$$

By (2), (3), and the choice of ψ in (4) this would entail

$$W\phi_k \geqq \min\left[W\phi_{k+1}, W(\phi_k - \phi_{k+1})\right] > \min\left[V_k\phi_{k+1}, V_k\phi_k\right] = V_k\phi_k = W\phi_k,$$

a contradiction which establishes the desired condition.

The inductive value V_{k+1} thus constructed satisfies the analogues of the desired conditions (1), (2), and (3). The latter two are consequences of the definitions in (4) and (7), while (1) follows from the definition (see §4, (3)) of the augmented value V_{k+1} by the triangle axiom for W:

$$W\left(\sum_{i=0}^{m} f_i(x)\psi^i\right) \geqq \min_i\left[Wf_i(x) + i\mu_{k+1}\right] = V_{k+1}\left(\sum_{i=0}^{m} f_i(x)\psi^i\right).$$

The inductive construction of the value V_k associated with W is complete.

This process either will ultimately yield an inductive value V_k equal to W or will give an infinite sequence of inductive values with a limit-value V_∞ such that

$$Wf(x) \geqq V_\infty f(x) = \lim_{k\to\infty} V_k f(x) \qquad \text{(for all } f(x)\text{)}.$$

In the discrete case the first inequality sign never occurs. For suppose instead that it did hold for some $f(x)$; then since $\{V_k f\}$ is monotone non-decreasing,

$$Wf(x) > V_k f(x) \qquad (k = 1, 2, \cdots).$$

The equivalence of (5) and (6) then implies that $f(x)$ is equivalence-divisible by $\phi_{k+1}(x)$ in V_k. Hence the monotonicity Theorem 5.1 shows

$$V_{k+1}f(x) > V_k f(x) \qquad (k = 1, 2, \cdots).$$

This cannot hold if the degrees of the key polynomials $\phi_k(x)$ increase indefinitely, so that we have the case where $\phi_k(x)$ has the fixed degree M for $k>t$, as in Theorem 7.1. The monotonic increasing sequence $\{V_k f(x)\}$ consists of numbers all from the discrete group Γ_t, with the result

$$Wf(x) \geqq V_\infty f(x) = \lim_{k\to\infty} V_k f(x) = \infty.$$

This can occur only for $f(x)=0$, a trivial case. Accordingly, $W=V_\infty$, and the completeness theorem is established.

Ostrowski ([18], pp. 361–392) has developed another method for finding all non-archimedean values of a transcendental extension $K(x)$ of a field K. His method makes no discreteness assumptions, but requires that K first be extended to an algebraically closed field A. The values of $A(x)$ and thereby those of $K(x)$ are constructed by means of "pseudo-convergent sequences" analogous to limit-values with linear key polynomials. A higher degree would be impossible by the character of A. Note, however, that his method requires an elaborate construction to obtain values of A from those of K, and that the use of A precludes any application to irreducibility criteria or to the algebraic extensions of the residue-class field discussed in our Part II.

II. The structure of inductive values

9. **Properties of key polynomials.** To apply the preceding construction of values to any particular case it is necessary to know what polynomials can be used as key polynomials. This question is not constructively answered by the definition in §4. Part of this question will be answered at once (Theorem 9.4); the rest after the structure of the inductive values V_k has been more explicitly formulated. We first show that certain polynomials act like "equivalence-units":

LEMMA 9.1. *If V_k is an inductive value with $k>1$, then for every polynomial $b(x)$ with $V_k b(x) = V_{k-1} b(x)$ there is a polynomial $b'(x)$ with*

$$(1) \qquad b'(x)b(x) \infty 1 \quad (\text{in } V_k), \qquad V_k b'(x) = V_{k-1} b'(x).$$

The hypothesis on $b(x)$ implies that $b(x)$ is not divisible by the last key polynomial $\phi_k(x)$. Since ϕ_k is certainly irreducible in the ordinary sense, there are polynomials $b'(x)$ and $c(x)$ with

$$b'(x)b(x) + c(x)\phi_k(x) = 1, \qquad \deg b'(x) < \deg \phi_k(x).$$

By Theorem 5.1, $V_k b' = V_{k-1} b'$. The transition from V_{k-1} to V_k increases the value of $c\phi_k$, but leaves unchanged the values of $b'b$ and 1 in this equation. Hence $b'b \infty 1$, as in (1).

LEMMA 9.2. *In any inductive V_k, the last key polynomial ϕ_k is equivalence-irreducible in V_k; a polynomial $g(x)$ not equivalence-divisible by ϕ_k in V_k has a value $V_k g$ in Γ_{k-1}.*

If a polynomial $f(x)$ has the expansion

$$(2) \qquad f(x) = f_n(x)\phi_k^n + f_{n-1}(x)\phi_k^{n-1} + \cdots + f_0(x), \qquad \deg f_i(x) < \deg \phi_k(x),$$

then $f(x)$ *is equivalence-divisible in V_k by ϕ_k if and only if $V_k f_0 > V_k f$.* For if $V_k f_0 > V_k f$, then $f - f_0$ is a polynomial equivalent to f with a factor ϕ_k. Con-

versely, if $f \infty h(x)\phi_k$, then the last term f_0 of the expansion for f is obtained from $f - h\phi_k$, where $V_k(f - h\phi_k) > Vf$, so that $V_k f_0 > V_k f$. In particular, an f not equivalence-divisible by ϕ_k has $V_k f = V_k f_0 = V_{k-1} f_0 \in \Gamma_{k-1}$, as asserted.

This criterion shows ϕ_k equivalence-irreducible in V_k. For suppose instead that $f(x)g(x)$ is equivalence-divisible by ϕ_k, although neither factor is so divisible. Then the criterion gives $V_k f_0 = V_k f$, $V_k g_0 = V_k g$, where $g_0(x)$ is the last term in the expansion for g. The last term in the expansion for fg is the remainder $r_0(x)$ obtained by dividing $f_0 g_0$ by ϕ_k; but since $f_0 g_0$ is not equivalence-divisible in V_{k-1} by ϕ_k, Lemma 4.3 shows

$$V_k r_0 = V_{k-1} r_0 = V_{k-1}(f_0 g_0) = V_k f_0 + V_k g_0 = V_k(fg).$$

This means that fg is not equivalence-divisible by ϕ_k, a contradiction proving the lemma.

An inductive value V_k will be called *commensurable* if the value μ_k assigned the last key polynomial is commensurable with the previous value-group Γ_{k-1} (cf. Theorem 6.7). There is then a smallest positive integer τ_k such that $\tau_k \mu_k$ is in Γ_{k-1}. For each $t \leqq k$ there is a similar τ_t:

(3) τ_t is the smallest integer such that $\tau_t \mu_t \in \Gamma_{t-1}$.

We will subsequently need polynomials with any given values:

LEMMA 9.3. *If V_k is a commensurable inductive value, then for any real number λ in the value-group Γ_k of V_k there is a polynomial $R_\lambda = R_\lambda(x)$ with value λ in V_k and in every value V_{k+1} obtained by augmenting V_k.*

Proof. As in Theorem 6.6, λ has the form

$$\lambda = \nu + m_1\mu_1 + m_2\mu_2 + \cdots + m_k\mu_k, \qquad \nu \in \Gamma_0.$$

Each integer m_i may be made non-negative by adding to $m_i\mu_i$ and subtracting from ν a sufficiently large term $q\mu_i$, so chosen that $q\mu_i \in \Gamma_0$ (e.g., choose $q \equiv 0 \pmod{\tau_1, \tau_2, \cdots, \tau_i}$). If then a is a constant of value ν,

$$R_\lambda = R_\lambda(x) = a x^{m_1} \phi_2^{m_2} \phi_3^{m_3} \cdots \phi_k^{m_k}, \qquad V_k R_\lambda = \lambda, \qquad m_i \geqq 0,$$

is the required polynomial. In any augmented value V_{k+1}, R_λ has value λ by Theorem 6.4.

THEOREM 9.4. *A polynomial $f(x)$ is a key polynomial for an inductive value over V_k if and only if the following conditions hold*: (i) *the expansion* (2) *has a last term with $V_k f = V_k f_0$*; (ii) *the expansion has a first term $f_n(x)\phi_k^n$ with $f_n(x) = 1$, $V_k \phi_k^n = V_k f$, and $n \equiv 0 \pmod{\tau_k}$*; (iii) *$f(x)$ is equivalence-irreducible in V_k.*

Proof. Condition (i) means, as in the proof of Lemma 9.2, that $f(x)$ is not equivalence-divisible in V_k by ϕ_k. Assume first that $f(x)$ is a key. Condition

(iii) is necessary by definition. Were (i) false, then $V_k f < V_k f_0$, so that $f \infty (f - f_0)$, while (2) shows $f - f_0 = q(x)\phi_k$ for a $q(x)$ of degree less than $f(x)$. Thus $f \infty q\phi_k$ in V_k. Since f is a key, this leads to a contradiction much as in the proof of Theorem 6.7. The assumption $V_k f_0 \neq V_k f$ is false.

Since $f(x)$ is minimal (Definition 4.1) it has no equivalence-multiples of degree less than itself. Hence $f_n(x)$ is a constant in K, for otherwise $k > 1$, and Lemma 9.1 supplies a $b'(x)$ with $b'(x)f_n(x) \infty 1$ in V_k. The product $b'(x)f(x)$ formed from (2) and modified by replacing the first coefficient by 1 and by reducing the other coefficients modulo ϕ_k is then an equivalence-multiple of $f(x)$. Its degree is $n \cdot \deg \phi_k$, and is less than that of $f(x)$ unless $f_n(x) \in K$. As the leading coefficient of f must be 1, $f_n(x) = 1$ follows, as in (ii). Certainly $V_k \phi_k^n = V_k f$ is necessary, for otherwise $f - \phi_k^n$ is an equivalent polynomial of smaller degree. Thus

$$V_k \phi_k^n = V_k f = V_k f_0 = V_{k-1} f_0 \in \Gamma_{k-1},$$

so that $n \equiv 0 \pmod{\tau_k}$ by (3). This establishes the necessity of (ii).

Conversely, if $f(x)$ satisfies (i), (ii), and (iii), it has first coefficient 1 and is minimal, because any equivalence-multiple of $f(x)$ must be of degree at least n in ϕ_k (cf. the proof of Theorem 4.2). The remaining restrictions of Definition 6.1 are readily verified, so that $f(x)$ is in fact a key polynomial.

10. **Residue-class fields.** The structure of a ring S with a value V involves the corresponding *value-ring* S^+, which consists of all elements a of S with $Va \geqq 0$ (these elements are the so-called "integers" of S). A *congruence* for integers can be defined thus

$$(1) \qquad a \equiv b \pmod{V} \text{ if and only if } V(a - b) > 0.$$

All elements of S^+ congruent to a given b form a *residue-class*; these classes together yield as usual the *residue-class ring* of V in S. This ring can also be considered as the residue-class ring S^+/P, where P, the set of all elements of S^+ with positive value, is a prime ideal in S^+. If S is a field, then S^+/P is also a field, the *residue-class field* of V in S. The structure of V depends essentially* on this residue-class field. For the p-adic value V_0 of the rational numbers (see §1, (1)) this field is simply the field of integers modulo p. Our problem is the determination of the residue-class field for any discrete inductive value.

If the residue-class of each integer a be denoted by Ha, then H is a homomorphism of S^+ to the residue-class ring $\Delta = S^+/P$, so that H has the following properties:

I. H is a many-one correspondence between S^+ and Δ;

* Hasse-Schmidt [12], p. 7; Ostrowski [18], p. 321.

II. H leaves sums and products unchanged; i.e., for $Va \geqq 0$ and $Vb \geqq 0$,

$$H(a+b) = Ha + Hb; \qquad H(ab) = (Ha)(Hb). \tag{2}$$

III. If $Va \geqq 0$, then $Ha = 0$ if and only if $Va > 0$.

By II, the last condition means that H carries congruent elements and only congruent elements into the same residue-class.

For an inductive value V_k we denote the residue-class rings thus, for $t = 1, 2, \cdots, k$:

(3) Λ_t is the residue-class field of V_t in $K(x)$;

(4) H_t is the homomorphism from $K(x)^+$ to Λ_t;

(5) Δ_t is the residue-class ring of V_t in $K[x]$.

But $f(x)$ and $g(x)$ are congruent as polynomials (mod V_t) if and only if they are congruent as rational functions (mod V_t). Hence each residue-class of Δ_t is contained in a residue-class of Λ_t, and no two residue-classes of Δ_t are contained in the same class of Λ_t. Addition and multiplication of classes are defined as addition and multiplication on elements in the classes, and hence are the same in Δ_t as in Λ_t. Therefore Δ_t is isomorphic to a subring of Λ_t. Since isomorphism does not alter the structure of a ring, we will *replace Δ_t henceforth by the isomorphic subring of Λ_t*. Then the H_t of (4) is also the homomorphism from $K[x]^+$ to Δ_t.

The correspondence H_t for rational functions is usually determined by the H_t for polynomials. For if $f(x)/g(x) \neq 0$ is a rational function with non-negative value and if V_t is commensurable, there is by Lemma 9.3 a polynomial $R(x)$ with $V_t R = -V_t g$, and by (2)

$$H_t\left(\frac{f(x)}{g(x)}\right) = H_t\left(\frac{Rf}{Rg}\right) = \frac{H_t(Rf)}{H_t(Rg)}. \tag{6}$$

Both $H_t(Rf)$ and $H_t(Rg)$ are residue-classes of polynomials, while $H_t(Rg) \neq 0$ by Property III. We have proved

LEMMA 10.1. *For a commensurable V_k, the residue-class field Λ_k of $K(x)$ is the quotient-field of the residue-class ring Δ_k of $K[x]$.*

THEOREM 10.2. *For a commensurable first stage inductive value $V_1 = [V_0, V_1 x = \mu_1]$, the residue-class ring Δ_1 is isomorphic to the ring $F_0[y]$ of all polynomials in a variable y with coefficients in F_0, the residue-class field of the value V_0 for K.*

Proof. There is given a homomorphism H_0 from the ring K^+ of all V_0-integers b in K to the residue-class field F_0. Each residue-class $H_1 b$ of Δ_1 contains the residue-class $H_0 b$ of F_0, and this correspondence $H_1 b \longleftrightarrow H_0 b$ is an isomorphism between F_0 and the set of those classes of Δ_1 containing elements

of K. We will identify F_0 with this isomorphic subfield of Δ_1; then Δ_1 is an extension of F_0 and $H_1 b = H_0 b$ for all b in K^+.

Any monomial bx^n of value zero has $V_0 b = -nV_1 x = -n\mu_1$, so that the exponent n is a multiple of the integer τ_1, defined in §9, (3). Any $f(x)$ with $V_1 f \geqq 0$ thus has the form

$$f(x) \equiv b_m x^{m\tau_1} + b_{m-1}x^{(m-1)\tau_1} + \cdots + b_1 x^{\tau_1} + b_0 \pmod{V_1}$$

after terms of positive value are omitted. If e is a constant in K of value $V_0 e = \tau_1\mu_1$, each term $b_j x^{j\tau_1}$ may be rewritten as a product $(b_j e^j)(e^{-j}x^{j\tau_1})$ of two factors of value 0. The application of the homomorphism H_1 then yields

$$H_1 f(x) = \sum_{j=0}^{m} H_0(b_j e^j) y^j; \qquad y = H(e^{-1}x^{\tau_1}). \tag{7}$$

With y so defined, any $H_1 f$ in Δ_1 becomes a polynomial in y with coefficients $H_0(b_j e^j)$ in F_0, so that the residue-class ring Δ_1 is contained in $F_0[y]$. Since $y \in \Delta_1$ and Δ_1 is a ring, $\Delta_1 = F_0[y]$. The element y is transcendental (i.e., a variable) over F_0, for otherwise it would satisfy an algebraic relationship $\alpha(y) = 0$, where

$$\alpha(y) = \alpha_m y^m + \alpha_{m-1}y^{m-1} + \cdots + \alpha_0; \qquad \alpha_m \neq 0, \qquad \alpha_j \in F_0.$$

Then the residue-class $\alpha(y)$ contains the polynomial

$$f(x) = a_m e^{-m} x^{m\tau_1} + a_{m-1}e^{-m+1}x^{(m-1)\tau_1} + \cdots + a_0,$$

where each a_j is a constant with $H_0 a_j = \alpha_j$. Then $V_1 f \geqq 0$ and $H_1 f = \alpha(y) = 0$, so that, by Property III of H_1, $V_1 f > 0$. But $\alpha_m \neq 0$, so that $V_0 a_m = 0$ and $V_1(a_m e^{-m}x^{m\tau_1}) = V_1 f = 0$, a contradiction. The theorem is established. We note also that (7) enables us to calculate the residue-class of any given $f(x)$.

11. **Conditions for equivalence-irreducibility.** A key polynomial ϕ_{k+1} over a value V_k is not equivalence-divisible by ϕ_k (Theorem 9.4, condition (i)). For any $f(x)$ with this property, questions of equivalence-divisibility can be handled as follows:

LEMMA 11.1. *In a commensurable V_k let $f(x)$ be a polynomial not equivalence-divisible by ϕ_k, and choose a polynomial $R(x)$ so that $V_{k-1}R = V_k R = -V_k f$. Then a polynomial $g(x)$ with $V_k g = 0$ is equivalence-divisible by f in V_k if and only if $H_k g$ is divisible by $H_k[Rf]$ in the residue-class ring Δ_k.*

By Lemma 9.2, $V_k f$ is in Γ_{k-1}, so Lemma 9.3 yields the R desired, and $H_k[Rf] \neq 0$. Suppose first that $H_k g$ (which is not 0 by Property III of H_k) is divisible by $H_k[Rf]$. Then $H_k g = \alpha \cdot H_k[Rf]$ for some residue-class $\alpha = H_k h(x) \neq 0$ in Δ_k, and

$$H_k g = \alpha \cdot H_k[Rf] = (H_k h)H_k[Rf] = H_k(hRf).$$

Thus g and hRf have the same residue-class, $V_k(g-hRf)>0=V_kg$, and $g\infty hRf$ is equivalence-divisible by f, as asserted.

Conversely, if g is equivalence-divisible by f, then $g\infty hf\infty hR'Rf$, where $R'(x)$ is a polynomial chosen as in Lemma 9.1 so that $RR'\infty 1$. But Rf, g, and hence hR' have value 0, so that

$$g \equiv hR'Rf \quad (\text{in } V_k); \qquad H_kg = H_k(hR')\cdot H_k[Rf],$$

which shows H_kg divisible by $H_k[Rf]$.

LEMMA 11.2. *For f and R as in Lemma* 11.1, *$f(x)$ is equivalence-irreducible in V_k if and only if every product in Δ_k divisible by $H_k[Rf]$ has* a factor divisible by $H_k[Rf]$ in Δ_k.*

Suppose first that f is equivalence-irreducible and that $(H_kg)(H_kh)=H_k(gh)$ is a multiple of $H_k[Rf]$. As we can assume $V_kg=V_kh=0$, the previous lemma shows the product gh equivalence-divisible by the equivalence-irreducible f, so that one of the factors is so divisible. By Lemma 11.1 this means that H_kg or H_kh is a multiple of $H_k[Rf]$, as asserted in the lemma.

Conversely, suppose that every product $(H_kg)(H_kh)$ divisible by $H_k[Rf]$ has a factor so divisible, and consider a product $a(x)b(x)$ equivalence-divisible by f, so that $ab\infty c(x)f$ for some c. Write $a(x)\infty g(x)\phi_k{}^d$ and $b(x)\infty h(x)\phi_k{}^e$, where the powers d and e are chosen so large that g and h are not equivalence-divisible by ϕ_k in V_k. Then V_kg and V_kh are by Lemma 9.2 in Γ_{k-1}, so that there are polynomials $S(x)$ and $T(x)$ with $V_k(gS)=V_k(hT)=0$. Then

$$STab \infty (Sg)(Th)\phi_k{}^{d+e} \infty STcf \qquad (\text{in } V_k).$$

But f is not equivalence-divisible by ϕ_k while ϕ_k is equivalence-irreducible (Lemma 9.2), so STc is equivalence-divisible by $\phi_k{}^{d+e}$. Removal of this factor makes $(Sg)(Th)$ equivalence-divisible by f, so that as in the previous lemma $H_k(Sg)H_k(Th)$ is divisible by $H_k[Rf]$. One of the factors, say $H_k(Sg)$, is then divisible by $H_k[Rf]$, and (Lemma 11.1) Sg is equivalence-divisible by f. But $a\infty S'(Sg)\phi_k{}^d$, where S' is chosen so that $S'S\infty 1$. Hence $a(x)$ is equivalence-divisible by f, and f is equivalence-irreducible.

12. Residue-class rings for commensurable values. We have

THEOREM 12.1. *If V_k is a commensurable inductive value of $K[x]$, given as in* §6, (3), *and if the original value V_0 of K has a residue-class field F_0, then there is a sequence of fields $F_1=F_0, F_2, F_3, \cdots, F_k$, each an algebraic extension of the preceding, such that for any $t=1, \cdots, k$ the V_t-residue-class ring of $K[x]$ is (isomorphic to) the ring $F_t[y]$ of polynomials in a variable y with coefficients in F_t. For $t>1$ the degree m_t of F_t is determined by* (cf. §9, (3))

* That is, the principal ideal $(H_k[Rf])$ is a prime ideal in Δ_k.

$$m_t\tau_{t-1}(\deg \phi_{t-1}) = \deg \phi_t; \qquad m_t = \deg [F_t : F_{t-1}].$$

By Lemma 10.1 we can then conclude at once

COROLLARY 12.2. *$F_t(y)$ is the V_t-residue-class field of $K(x)$.*

The case $t=1$ of this theorem is known (Theorem 10.2); hence we use induction, and assume the theorem true for V_t. It is convenient to omit the subscript $t+1$ (but not the subscript t) and to write V, ϕ, H, τ, etc., for V_{t+1}, ϕ_{t+1}, H_{t+1}, τ_{t+1}, etc. By the monotonic character of V (Theorem 5.1) polynomials $f(x)$ and $g(x)$ with $V_t f \geqq 0$ and $V_t g \geqq 0$ are congruent mod V_t only if they are congruent mod V. Each residue-class mod V_t is thus contained in a residue-class mod V, and this gives a homomorphism between $\Delta_t = F_t[y]$ and a subring F of the residue-class ring Δ (cf. §10, (5)), where $F = F_{t+1}$ *is composed of all residue-classes* mod V *containing an* $f(x)$ *with* $V_t f \geqq 0$. Polynomials f and g incongruent mod V_t become congruent mod V if and only if $f-g$ is equivalence-divisible by ϕ in V_t (Theorem 5.1). This means that $H_t f - H_t g$ is divisible by the polynomial

$$(1) \qquad \psi_{t+1}(y) = \psi(y) = H_t[R\phi]; \qquad (V_{t-1}R = V_tR = -V_t\phi,\ R = R_{t+1}(x)),$$

constructed as in Lemma 11.1. Since not all polynomials are equivalence-divisible by ϕ in V_t, $\psi(y)$ is not a constant in F_t, while Lemma 11.2 shows $\psi(y)$ an irreducible polynomial in $F_t[y]$. In the above homomorphism between $F_t[y]$ and F the multiples of $\psi(y)$ in $F_t[y]$ are the elements corresponding to 0, so that F is isomorphic to the ring of polynomials $F_t[y]$ modulo $\psi(y)$, or, alternatively, to the field obtained by adjoining to F_t a root θ of $\psi(y)$. We identify F with this isomorphic field:

$$(2) \qquad F = F_{t+1} = F_t(\theta); \qquad \psi(\theta) = 0 \qquad (\theta = \theta_{t+1}).$$

Then the residue-class $H_t f$, when reduced modulo $\psi(y)$, will be identical to the residue-class Hf; that is,

$$(3) \qquad V_t f(x) \geqq 0 \text{ implies } Hf = [H_t f]_{y=\theta}.$$

A monomial expansion $a(x)\phi^n$ of value 0 must have n a multiple of τ (cf. §9, (3)). Hence any $f(x)$ with $Vf \geqq 0$ has the form

$$(4) \quad f(x) \equiv f_m(x)\phi^{m\tau} + f_{m-1}(x)\phi^{(m-1)\tau} + \cdots + f_0(x) \pmod V, \deg f_i < \deg \phi.$$

Since $V\phi^\tau \in \Gamma_t$, there are by Lemmas 9.1 and 9.3 polynomials $Q_{t+1}(x)$ and $Q'_{t+1}(x)$ such that

$$(5) \qquad V_tQ = VQ = V\phi^\tau, \qquad QQ' \equiv 1 \pmod V, \qquad V_tQ' = VQ' = -V\phi^\tau.$$

The terms $f_i\phi^{i\tau}$ in the expansion (4) can be rewritten as products $(f_iQ^i)(\phi^{i\tau}Q^{-i})$,

where $V_t(f_jQ^j) \geqq 0$, $V(\phi^{j\tau}Q^{-j}) = 0$, the former because $Vf \geqq 0$. The application of H, with (3), then proves

$$Hf(x) = \sum_{j=0}^{m} [H_t(f_j(x)Q^j)]_{y=\theta} \cdot y_1^j; \qquad y_1 = H(\phi^\tau Q^{-1}). \tag{6}$$

This shows that every Hf in the residue-class ring Δ is also in $F[y_1]$, while by (5), $y_1 = H(\phi^\tau Q^{-1}) = H(\phi^\tau Q')$ is a residue-class of a polynomial, hence is in Δ. Consequently, $F[y_1] = \Delta$, as asserted in the theorem.

The element y_1 is transcendental over F; for suppose instead that y_1 satisfied an algebraic relation $\alpha(y_1) = 0$, with

$$\alpha(y_1) = \alpha_m y^m + \alpha_{m-1} y^{m-1} + \cdots + \alpha_0; \qquad \alpha_m \neq 0, \qquad \alpha_j \in F.$$

By the original (italicized) definition of F each residue-class α_j of F contains a polynomial $b_j(x)$ with $V_t b_j \geqq 0$, so that $Hb_j = \alpha_j$. Then

$$f(x) = \sum_{j=0}^{m} b_j(x) Q'^j \phi^{j\tau} \equiv \sum_{j=0}^{m} b_j(x) Q^{-j} \phi^{j\tau} \qquad \text{(in } V\text{)}$$

is a polynomial of non-negative value which has $Hf = \alpha(y_1) = 0$. By Property III of H, $Vf > 0$. On the other hand Vf must equal 0, for $Hb_m = \alpha_m \neq 0$ gives $Vb_m = 0$ and $Vf \leqq V(b_m Q^{-m} \phi^{m\tau}) = 0$, by Theorem 5.2. This contradiction shows y_1 a variable over F.

The formula (6) enables us to calculate $Hf(x)$ effectively for any $f(x)$ given in (4), provided only that $V_t(f_jQ^j) \geqq 0$ for all j.

It remains to determine the degree of the field F over F_t, which by (2) is the degree of $\psi(y)$. The key ϕ has by Theorem 9.4 an expansion of the form

$$\phi = \phi_t^{m\tau_t} + \sum_{i=0}^{m\tau_t - 1} a_i(x) \phi_t^i, \qquad V_t\phi = V_t\phi^{m\tau_t} = V_t a_0. \tag{7}$$

If $t > 1$, $\psi = H_t[R\phi]$ can be computed by the analog of (6) for the preceding stage (with t in (6) replaced by $t-1$), for the coefficients Ra_i must by the choice of R have $V_{t-1}(Ra_i) \geqq 0$. This calculation shows $\psi(y)$ to be a polynomial in y with a first term $H_t(RQ_t^m)y^m$ arising from the first term of (7). But

$$V_t(RQ_t^m) = V_tR - V_tQ_t^m = V_t\phi - V_t\phi_t^{m\tau_t} = 0,$$

so that the coefficient of y^m is not 0. The polynomial ψ has degree m, and by (7)

$$m\tau_t \deg \phi_t = \deg \phi, \qquad m = \deg \psi = \deg [F:F_t],$$

as asserted* in Theorem 12.1. This theorem has now been demonstrated.

* The proof given holds for $t > 1$, but may be simplified for the case $t = 1$.

13. **Conditions for key polynomials.** In the criterion of Theorem 9.4 for a key polynomial the condition (iii) of equivalence-irreducibility can now be replaced by the condition of Lemma 11.2,

(iiia), $H_k[Rf]$ is an irreducible polynomial in $F_k[y]$.

This yields a final explicit description of key polynomials. A partial converse is possib'e:

THEOREM 13.1. *In a given V_k, let $\psi(y) \neq y$ be a polynomial* of degree $m>0$, irreducible in $F_k[y]$ and with first coefficient 1. Then there is one and, except for equivalent polynomials in V_k, only one $\phi(x)$ which is a key polynomial and which has $H_k[R\phi]=\psi(y)$, for a suitable R chosen as in* §11.

Proof. There is a polynomial $f(x)$ with the residue-class ψ, so that $H_kf=\psi$ and $V_kf=0$. If we multiply f by Q_k^m (chosen as in §12, (5)) and in the expansion of the resulting product drop all terms not of minimum value and then replace the leading coefficient of ϕ_k by 1, we obtain a polynomial $\phi(x)$ with the value $V_kQ_k^m$. For R we can then use $Q_k'^m$, so that

$$H_k[R\phi] = H_k[Q_k'^mQ_k^mf] = H_kf = \psi(y).$$

Furthermore ϕ can be shown to satisfy the remaining conditions (i) and (ii) of Theorem 9.4, hence ϕ is a key polynomial. The uniqueness is readiiy established.

Since $H_k[Rf]$ can be effectively constructed by §12, (6), the problem of testing whether a given $f(x)$ is a key polynomial is reduced to that of testing the image $H_k[Rf]$ of $f(x)$ for irreducibility in $F_k[y]$. If K is the field of rationals, then F_k is a finite field and the latter problem is completely solvable. This result can be used to construct examples for inductive values of any stage and for limit-values of both constant degree† and increasing degree types. The construction of constant degree values may be simplified by deducing from Theorem 9.4 the following partial converse of Lemma 6.3:

COROLLARY 13.2. *In V_k let $s(x)$ be a polynomial of degree less than that of $\phi_k(x)$ and with $V_ks(x)=V_k\phi_k$. Then $\phi_k(x)+s(x)$ is a key polynomial for an inductive value over V_k.*

14. **Special cases of homomorphism.** The residue-class fields can be similarly found for finite discrete limit-values and for inductive values where the value for K is trivial (§2) or where the last assigned value μ_k is incommensurable (§6).

* The assumption $\psi(y)\neq y$ is needed, for the condition $V_kf=V_kf_0$ in Theorem 9.4 implies $H_k[Rf]\neq y$.

† By using a transcendental p-adic number the finiteness condition of §7 can be satisfied.

THEOREM 14.1. *Let V_∞ be a limit-value constructed as in Theorem 6.1 from a sequence of values $V_0, V_1, V_2, \cdots$, and satisfying one of the conditions*:

(a) *the degrees of the keys ϕ_k are not bounded as $k \to \infty$;*

(b) *V_∞ is finite and discrete, and* $\deg \phi_k = M$ *for all $k > t$.*

The fields F_k of Theorem 12.1 yield a (possibly infinite) extension field $F_\infty = F_0 + F_1 + F_2 + \cdots$ which is isomorphic to the V_∞ residue-class field both for $K[x]$ and for $K(x)$. In case (b), $F_\infty = F_{t+1}$.

Let H_k be the homomorphism of $K[x]$ to the V_k residue-class ring $F_k[y]$ and H_∞ the homomorphism of $K[x]$ to the V_∞ residue-class ring Δ_∞. Then

$$(1) \qquad V_{k-1}f(x) \geqq 0 \text{ implies that } H_{k+j}f = H_k f \in F_k \qquad (j = 1, 2, \cdots).$$

For, according to §12, (3),

$$H_k f = [H_{k-1}f]_{y=\theta_k}, \qquad H_{k+1}f = [H_k f]_{y=\theta_{k+1}},$$

which indicates that $H_k f$ is a constant free of y in F_k, and that $H_{k+1}f$ must equal $H_k f$, and so on.

In case (a) there is for every $f(x)$ a k so large that $\deg \phi_k > \deg f$, so that $V_\infty f = V_{k-1}f$, as in Theorem 5.1. If $V_\infty f \geqq 0$, then, by (1), $H_{k+j}f = H_k f$ is a constant in F_k independent of j. The correspondence

$$(2) \qquad H_\infty f \longleftrightarrow H_k f, \text{ for } k \text{ with } H_k f = H_{k+1}f = H_{k+2}f = \cdots,$$

carries each element of Δ_∞ into an element of F_∞. Every element α of F_∞ is used, for, by the definition of F_∞, α is in some F_k so that α has the form $H_k f$, and $H_{k+j}f = H_k f$ as in (1), whence α corresponds in (2) to $H_k f$. The correspondence (2) is one-one, for elements are congruent mod V_∞ if and only if they are congruent modulo some V_k. Finally, (2) is an isomorphism, making $F_\infty \sim$ or $= \Delta_\infty$ as asserted. The residue-class field of $K(x)$ is, by the argument of Lemmas 9.3 and 10.1, just the quotient field of F_∞, and must then be F_∞ itself.

In the case (b), the degrees of the extensions $F_{k+1} : F_k$ as determined in Theorem 12.1 are all 1 for $k > t$. Hence $F_k = F_{t+1}$. Because V_∞ is finite (§7) and discrete, Theorem 6.5 yields for any $f(x)$ with $V_\infty f \geqq 0$ an $i \geqq t$ so large that $V_\infty f = V_i f \geqq 0$. Then $H_k f$ is again ultimately constant, and (2) gives the isomorphism as before.

THEOREM 14.2. *For an incommensurable inductive value V_k of $K(x)$ the field F_k determined from F_{k-1} and ϕ_k exactly as in §12 is the V_k residue-class field of both $K[x]$ and $K(x)$.*

Proof. Since no non-zero multiple of $\mu_k = V_k \phi_k$ lies in Γ_{k-1}, no two terms in a ϕ_k-expansion can have the same value in V_k. Hence any polynomial is

equivalent to a monomial expansion, and every rational function has by Lemma 9.1 the form

$$f(x)/g(x) \sim c(x)\phi_k^m \quad (\text{in } V_k), \qquad V_k c = V_{k-1} c.$$

If f/g has value 0, then $m=0$, $V_{k-1}c \geqq 0$, and $H_k(f/g) = H_k c$. But F_k is defined in §12, italics (or, for $k=1$, in §10) as all residue-classes $H_k h$ with $V_{k-1}h \geqq 0$. In this case every residue-class has this form, so that F_k is as asserted the whole residue-class field, either for $K[x]$ or for $K(x)$.

In particular, over the trivial value V_0(§2) of K the only non-trivial inductive values are

$$V_1 = [V_0, V_1 x = \mu_1], \qquad \mu_1 \neq 0;$$
$$V_2 = [V_0, V_1 x = 0, V_2 \phi = \mu_2], \qquad \mu_2 > 0, \phi(x) \text{ irreducible}.$$

Both are incommensurable (no multiple of μ_2 lies in the group Γ_0, which contains only 0). Furthermore, the residue-class field of K for the trivial V_0 is K itself. Hence the residue-class field for V_1 is K and for V_2 is $K(\theta)$, where θ is a root of $\phi(x)$.

15. Equality conditions for values. An inductive value is essentially a representation; the same value of $K[x]$ could easily have several such representations. This section and the next one will formulate necessary and sufficient conditions for the equality of two inductive or limit-values. In this connection two values V and W of a ring S will be called equal if and only if

$$Va = Wa \qquad (\text{all } a \in S). \tag{1}$$

In this section we consider the case when each key polynomial ϕ_k exceeds the preceding ϕ_{k-1} in degree—a case which can often be made to apply by omitting any ϕ_{k-1} without the above property.

LEMMA 15.1. *If an inductive value*

$$V_k = [V_{k-2}, V_{k-1}\phi_{k-1} = \mu_{k-1}, V_k \phi_k = \mu_k] \qquad (k \geqq 2)$$

has two key polynomials $\phi_{k-1}(x)$ and $\phi_k(x)$ of the same degree, then

$$W = [V_{k-2}, W\phi_k = \mu_k]$$

is an inductive value equal to V_k.

We first prove W an inductive value. Since ϕ_k exceeds $\phi_k - \phi_{k-1}$ in degree, the constant-degree Lemma 6.3 shows that

$$V_{k-2}(\phi_k - \phi_{k-1}) = V_{k-1}(\phi_k - \phi_{k-1}) = \mu_{k-1}, \mu_k > \mu_{k-1} > V_{k-2}\phi_{k-1}. \tag{2}$$

A combination of these two results proves

$$\phi_k \sim \phi_{k-1} \qquad (\text{in } V_{k-2}). \tag{3}$$

Thus ϕ_k and ϕ_{k-1} have the same equivalence-divisibility properties in V_{k-2}, and so ϕ_k, like ϕ_{k-1}, is a key polynomial over V_{k-2}. By (2) and (3) the value $\mu_k > V_{k-2}\phi_k$ assigned to ϕ_k is sufficiently large. Therefore W is inductive, for conditions (1) and (2) of Definition 6.1 hold trivially. The definition of augmented values applied to the usual expansion (e.g., §9, (2)) of any $f(x)$ in powers of ϕ_k gives

$$V_k f(x) = \min_i [V_{k-1} f_i(x) + i\mu_k], \qquad Wf(x) = \min_i [V_{k-2} f_i(x) + i\mu_k].$$

The corresponding terms $V_{k-1}f_i$ and $V_{k-2}f_i$ are equal by Theorem 5.1, for each $f_i(x)$ has a degree less than that of ϕ_k or of ϕ_{k-1}. Therefore $V_k = W$. Successive applications of this lemma give

THEOREM 15.2. *Any inductive value is equal to an inductive value in which* $\{\deg \phi_k\}$ *is a monotone increasing sequence. A similar representation holds for any limit-value not of constant-degree type.*

For values in this particular form we can obtain necessary and sufficient conditions for equality.

THEOREM 15.3. *If the two inductive values*

$$V_s = [V_0,\ V_1 x = \mu_1,\ V_2\phi_2 = \mu_2, \cdots,\ V_s\phi_s = \mu_s] \tag{4}$$

$$W_t = [W_0, W_1 x = \nu_1,\ W_2\psi_2 = \nu_2, \cdots, W_t\psi_t = \nu_t] \tag{5}$$

both have a monotone character such that

$$1 < \deg \phi_2 < \cdots < \deg \phi_s, \qquad 1 < \deg \psi_2 < \cdots < \deg \psi_t,$$

then $V_s = W_t$ *holds if and only if*

(i) $V_0 = W_0, \quad s = t;$

(ii) $\deg \phi_k(x) = \deg \psi_k(x) \qquad (k = 1, \cdots, t);$

(iii) $V_{k-1}(\psi_k - \phi_k) \geqq \mu_k = \nu_k \qquad (k = 1, \cdots, t).$

The theorem is still true if either s or t is $+\infty$.

First prove the sufficiency of these conditions. Since (i) and (iii) make V_1 and W_1 identical, we can proceed by induction, assuming that $V_{k-1} = W_{k-1}$ is already established. Now compute the V_k value of ψ_k. Because key polynomials have the leading coefficient unity, (ii) shows that the degree of ϕ_k exceeds that of $\psi_k - \phi_k$, so that ψ_k has the expansion

$$\psi_k = \phi_k + (\psi_k - \phi_k) \tag{6}$$

in powers of ϕ_k. The definition of V_k and (iii) prove

$$V_k\psi_k = \min\ [\mu_k, V_{k-1}(\psi_k - \phi_k)] = \mu_k = \nu_k.$$

The V_k value of any polynomial $f(x)$ can now be estimated from the expansion of $f(x)$ in ψ_k, for the triangle axiom for V_k gives

$$V_k\left[\sum_{j=0}^{n} f_j\psi_k^j\right] \geqq \min_j \left[V_{k-1}f_j + j\cdot V_k\psi_k\right] = \min_j \left[V_{k-1}f_j + j\cdot\nu_k\right] = W_kf,$$

because of the definition of W_k. Thus $V_kf \geqq W_kf$, while the inverse inequality is similarly proved. Hence $V_k = W_k$, and the induction is complete.

The necessity of the conditions depends chiefly on the invariance of the values assigned the key polynomials (Theorem 6.4). The assumption $V_s = W_t$ shows that $V_0 = W_0$ and $\mu_1 = \nu_1$. Hence (ii) and (iii) hold for $k=1$. We prove them by induction on k. If they hold through $k-1$, then the sufficiency proof shows $V_{k-1} = W_{k-1}$. By Theorem 5.1, $\deg \phi_k$ can be characterized as the smallest degree of any polynomial $a(x)$ with the property that $V_sa > V_{k-1}a$. Since V_s and W_t are equal, $\deg \psi_k$ can be characterized by the same statement, so that

$$\deg \phi_k(x) = \deg \psi_k(x). \tag{7}$$

The monotonic assumption on $\{\deg \phi_k\}$ then shows $V_s\psi_k = V_k\psi_k$. Hence, because of the invariance in W of the value assigned to ψ_k,

$$\nu_k = W_t\psi_k = V_s\psi_k = V_k\psi_k = V_k[\phi_k + (\psi_k - \phi_k)].$$

As before, (6) is an expansion in powers of ϕ_k, so that this equation becomes

$$\nu_k = \min\left[\mu_k, V_{k-1}(\psi_k - \phi_k)\right].$$

Combining this with the symmetric conclusion (using $V_{k-1} = W_{k-1}$)

$$\mu_k = \min\left[\nu_k, V_{k-1}(\psi_k - \phi_k)\right],$$

we obtain (iii) for index k. With (7) this completes the induction. The condition $s=t$ results, even in the case $s=t=\infty$.

16. **Normal forms for values.** The results of the previous section do not apply to constant degree limit values, nor do they yield unique normal forms. Both these goals can be reached in the discrete case by using key polynomials from which all unnecessary high-valued terms have been dropped.

In the expansion of any $f(x)$ in a value V_k, the coefficient $f_i(x)$ of any power of ϕ_k can itself be expanded in powers of ϕ_{k-1}. Since the degree of $f_i(x)$ is limited, the highest power of ϕ_{k-1} occurring is less than n_k/n_{k-1}, where n_i has the meaning

$$n_i = \deg \phi_i(x) \qquad (i = 1, \cdots, k). \tag{1}$$

By an inductive process of this sort one can prove

THEOREM 16.1. *In any V_k every polynomial $f(x)$ can be expanded as a polynomial in the key polynomials with constant coefficients,*

$$f(x) = \sum_j a_j \phi_1^{m_{1j}} \phi_2^{m_{2j}} \cdots \phi_k^{m_{kj}}, \qquad (a_j \in K), \tag{2}$$

where the exponents m_{ij} are limited as follows

$$m_{ij} < n_{i+1}/n_i \quad (\text{all } j;\ i = 1, 2, \cdots, k-1). \tag{3}$$

The value of $f(x)$, when computed from the definition, is

$$V_k f(x) = \min_j V_k(a_j \phi_1^{m_{1j}} \cdots \phi_k^{m_{kj}}). \tag{4}$$

For a p-adic value, every number is equivalent to one of the numbers $c \cdot p^m$, $c=0, 1, \cdots, p-1$. For any value V_0 of a field K we can similarly (axiom of choice) pick from each class of equivalent elements a single representative element; in particular, we can make 1 one of the representatives. Given fixed representatives of this sort for each V_0, we say that a polynomial $f(x)$ is *homogeneous* in a value V_k derived from V_0 if in the expansion (2) of $f(x)$ all the coefficients a_j are representatives in V_0 and all the terms have the same minimum value $V_k f(x)$.

LEMMA 16.2. *Every polynomial $f(x)$ is equivalent in V_k to one and only one homogeneous polynomial $h(x)$. This $h(x)$ is called the "homogeneous part" of $f(x)$.*

Proof. Given $f(x)$, we find $h(x)$ by altering coefficients and dropping out terms in the expansion (2) for f. Were $f(x)$ also equivalent to a homogeneous $g(x)$, then all terms in the expansions of both $h(x)$ and $g(x)$ would have the same value $V_k h$, while $h-g$ would have a larger value. Thus corresponding coefficients are equivalent and therefore equal.

An inductive or limit value $V_k = [V_0, V_1\phi_1 = \mu_1]$ may be called *homogeneous* if every key polynomial $\phi_i(x)$ is homogeneous in V_{i-1} $(i=2, \cdots, k)$. We will prove

THEOREM 16.3. *Any inductive or limit-value constructed from a discrete value V_0 of K is equal to a homogeneous inductive or limit-value.*

We have to prove that, if $U = [V_k, U\phi = \mu]$ is an augmented value over a homogeneous value V_k, then U itself is equal to a homogeneous inductive value. This is done by introducing successive homogeneous parts of ϕ as new keys. First use $\psi_1(x)$, the homogeneous part of ϕ in V_k. By Lemma 16.2

$$\psi_1(x) \sim \phi(x) \quad (\text{in } V_k), \qquad \deg \psi_1(x) = \deg \phi(x). \tag{5}$$

It follows that $\psi_1(x)$ is a key polynomial over V_k. Setting

$$\nu_1 = V_k[\phi(x) - \psi_1(x)] > V_k\psi_1(x),$$

we can construct a homogeneous value $W_1=[V_k, W_1\psi_1=\nu_1]$. If $\mu\leqq\nu_1$, then the sufficiency proof of Theorem 15.3 shows $U=W_1$. Otherwise $\mu>\nu_1$, and Corollary 13.2 proves $W'=[W_1, W'\phi=\mu]$ an inductive value, which by Lemma 15.1 is equal to U. We repeat the above argument, constructing a W_2 from $\psi_2(x)$, the principal part of ϕ in W_1. This gives a sequence of homogeneous inductive values,

$$W_t = [V_k, W_1\psi_1 = \nu_1, W_2\psi_2 = \nu_2, \cdots, W_t\psi_t = \nu_t] \qquad (t = 1, 2, \cdots).$$

The degrees of the $\psi_i(x)$ are all identical by (5), so that Lemma 6.3 proves that

$$\nu_1 < \nu_2 < \nu_3 < \cdots \tag{6}$$

and that each ν_i is in the value-group Γ_k of V_k. By hypothesis and Theorem 6.7, this Γ_k is discrete. Hence there is a smallest t with $\nu_t\geqq\mu$ in (6), and U is equal to the homogeneous value W_t. The advantage of so representing every value in a homogeneous form lies in the following uniqueness theorem:

THEOREM 16.4. *Two homogeneous inductive or limit-values which are equal must be identical.*

If the equal values are V_s and W_t as in §15, (4) and (5), then the asserted identity means simply that

$$V_0 = W_0, \qquad s = t, \tag{7}$$

$$\phi_k = \psi_k, \qquad \mu_k = \nu_k \qquad (k = 1, 2, \cdots, s). \tag{8}$$

The hypotheses readily give $V_0=W_0$ and (8) for $k=1$. Suppose (8) true up to $k-1$ inclusive. Then $V_{k-1}=W_{k-1}$. We can assume $s>k-1$, whence also $t>k-1$. Then ϕ_k has the following invariant properties which refer only to $V_{k-1}=W_{k-1}$ and $V_s=W_t$: ϕ_k is totally homogeneous in V_{k-1}, it has the first coefficient 1 and it has the minimum degree consistent with the property $V_s\phi_k>V_{k-1}\phi_k$. Furthermore ψ_k has the same properties. But these properties uniquely determine ϕ_k, for, since the difference $\phi_k-\psi_k$ is of degree less than ϕ_k, its value is

$$V_{k-1}(\phi_k - \psi_k) = V_s(\phi_k - \psi_k) \geqq \min [V_s\phi_k, V_s\psi_k] > \min [V_{k-1}\phi_k, V_{k-1}\psi_k].$$

Hence by the triangle law $\phi_k\infty\psi_k$ in V_{k-1}, so that by Lemma 16.2, $\phi_k=\psi_k$. By Theorem 6.4, $\mu_k=\nu_k$, as in (8). The induction ends when k reaches $s=t$, and the identity $V_s\equiv W_t$ is proved. A simple consequence is

COROLLARY 16.5. *If every value of K is discrete, then no inductive value can ever equal a limit-value, and no limit-value of constant degree type* (§7) *can equal a limit-value not of this type.*

Bibliography

Papers on irreducibility:

1. M. Bauer, *Verallgemeinerung eines Satzes von Schönemann*, Journal für die Mathematik, vol. 128 (1905), pp. 87–89.

2. H. Blumberg, *On the factorization of expressions of various types*, these Transactions, vol. 17 (1916), pp. 517–544.

3. G. Dumas, *Sur quelques cas d'irreductibilité des polynômes à coéfficients rationnels*, Journal de Mathématiques, (6), vol. 2 (1906), pp. 191–258.

4. G. Eisenstein, *Ueber die Irreduzibilität und einige andere Eigenschaften der Gleichungen*, etc., Journal für die Mathematik, vol. 39 (1850), p. 166.

5. L. Königsberger, *Ueber den Eisensteinschen Satz von der Irreduzibilität algebraischer Gleichungen*, Journal für die Mathematik, vol. 115 (1895), pp. 53–78; especially (67) on p. 69.

6. J. Kürschák, *Irreduzible Formen*, Journal für die Mathematik, vol. 152 (1923), pp. 180–191.

7. O. Ore, *Zur Theorie der Irreduzibilitätskriterien*, Mathematische Zeitschrift, vol. 18 (1923), pp. 278–288.

8. O. Ore, *Zur Theorie der Eisensteinschen Gleichungen*, Mathematische Zeitschrift, vol. 20 (1924), pp. 267–279.

9. O. Perron, *Neue Kriterien für die Irreduzibilität algebraischer Gleichungen*, Journal für die Mathematik, vol. 132 (1907), p. 304.

10. T. Rella, *Ordnungsbestimmungen in Integritätsbereichen und Newtonsche Polygone*, Journal für die Mathematik, vol. 158 (1927), pp. 33–48.

11. Th. Schönemann, *Von denjenigen Moduln, welche Potenzen von Primzahlen sind*, Journal für die Mathematik, vol. 32 (1846), pp. 93–105, §61.

Papers on absolute values:

12. H. Hasse and F. K. Schmidt, *Die Struktur diskret bewerteter Körper*, Journal für die Mathematik, vol. 170 (1933), pp. 4–63.

13. W. Krull, *Idealtheorie in unendlichen algebraischen Zahlkörpern*, II, Mathematische Zeitschrift, vol. 31 (1930), pp. 527–557.

14. W. Krull, *Allgemeine Bewertungstheorie*, Journal für die Mathematik, vol. 167 (1932), pp. 161–196.

15. J. Kürschák, *Limesbildung und allgemeine Körpertheorie*, Journal für die Mathematik, vol. 142 (1912), pp. 211–253.

16. S. MacLane, *Abstract absolute values which give new irreducibility criteria*, Proceedings of the National Academy of Sciences, vol. 21 (1935), pp. 272–274.

17. A. Ostrowski, *Ueber einige Lösungen der Funktionalgleichung* $\phi(x)\cdot\phi(y)=\phi(xy)$, Acta Mathematica, vol. 41 (1917), pp. 271–284.

18. A. Ostrowski, *Untersuchungen zur arithmetischen Theorie der Körper*, Mathematische Zeitschrift, vol. 39 (1934), pp. 269–404.

Harvard University,
Cambridge, Mass.

UNIFORMLY CONVEX SPACES*

BY

JAMES A. CLARKSON†

1. Introduction

The spaces with which we shall deal in this paper are Banach spaces, that is, linear, metric, complete, normed spaces, which possess in addition a certain property of *convexity* of the norm. Expressed in geometrical terms this property is simple: it is that the mid-point of a variable chord of the unit sphere of the space cannot approach the surface of the sphere unless the length of the chord goes to zero. Additional interest is given to the notion by the fact that, as we shall prove, the well known spaces L_p and l_p possess this property for p exceeding unity.

Several writers‡ have considered the problem of defining an integral of a function whose domain is in Euclidean space (or even a more general space) and whose range lies in a Banach space. Bochner§ has pointed out that such a function may be absolutely continuous without being an integral in his sense, or indeed without being differentiable at any point. We shall prove that if the range space is uniformly convex in our sense such phenomena do not occur, and that for these spaces the situation is quite analogous to the theory for ordinary complex functions.

2. Uniformly convex spaces

Let B denote a Banach space, with elements $x, y, \cdots$. We denote the norm of an element x by $\|x\|$.

Definition 1. *A Banach space B will be said to be uniformly convex if to each ϵ, $0<\epsilon\leqq 2$, there corresponds a $\delta(\epsilon)>0$ such that the conditions*

$$\|x\| = \|y\| = 1, \qquad \|x - y\| \geqq \epsilon$$

* Presented to the Society, February 29, 1936; received by the editors January 21, 1936.

† National Research Fellow.

‡ L. M. Graves, *Riemann integration and Taylor's theorem in general analysis*, these Transactions, vol. 29 (1927), pp. 163–172.

S. Bochner, *Integration von Funktionen, deren Werte die Elemente eines Vektorraumes sind*, Fundamenta Mathematicae, vol. 20 (1933), pp. 262–276.

N. Dunford, *Integration in general analysis*, these Transactions, vol. 37 (1935), pp. 441–453.

G. Birkhoff, *Integration of functions with values in a Banach space*, these Transactions, vol. 38 (1935), pp. 357–378.

§ See Bochner, *Absolut-additive abstracte Mengenfunktionen*, Fundamenta Mathematicae, vol. 21 (1933), pp. 211–213. We take this opportunity to acknowledge our indebtedness to Professor Bochner for his suggestions concerning this paper.

imply

$$\left\| \frac{x+y}{2} \right\| \leqq 1 - \delta(\epsilon).$$

We remark that Euclidean spaces of all dimensions, Hilbert space, and hyper-Hilbert spaces,* are all uniformly convex. This follows, for example, from the identity

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2) \tag{1}$$

which is known† to be characteristic of such spaces.

We shall need to recall the definition of the *product*‡ of a finite number of Banach spaces. Let $B_1, B_2, \cdots, B_k$ be k Banach spaces with elements $x^1, x^2, \cdots, x^k$. The product space $B = B_1 \times B_2 \times \cdots \times B_k$ is defined as the set of all ordered k-tuples $x = (x^1, x^2, \cdots, x^k)$, where addition and complex multiplication are defined in obvious fashion, and where the norm of the elements x of B is required, in addition to the usual properties, to have the property that $\|x\| \to 0$ is equivalent to $\|x^i\| \to 0$ $(i = 1, 2, \cdots, k)$. B is also a Banach space.

It is clear that the product of uniformly convex Banach spaces is not in general uniformly convex unless we require something more of the norm in the product space. In our first theorem we lay down a condition sufficient for this.

Let $N(a_1, a_2, \cdots, a_k)$ be a non-negative continuous function of the non-negative variables a_i. We say that N is

(a) *homogeneous*, if for $c \geqq 0$,

$$N(ca_1, ca_2, \cdots, ca_k) = cN(a_1, a_2, \cdots, a_k);$$

(b) *strictly convex*, if

$$N(a_1 + b_1, a_2 + b_2, \cdots, a_k + b_k) < N(a_1, a_2, \cdots, a_k) + N(b_1, b_2, \cdots, b_k)$$

unless $a_i = cb_i$ $(i = 1, 2, \cdots, k)$. In the latter case we have equality by condition (a);

(c) *strictly increasing*, if it is strictly increasing in each variable separately.

A familiar example of a function N satisfying these conditions is

* That is, those spaces which satisfy all the postulates of Hilbert space except that of separability. See J. v. Neumann, *Mathematische Grundlagen der Quantenmechanik*, Berlin, 1932, pp. 37–38, for discussion of an example.

† J. v. Neumann and Jordan, *On inner products in linear metric spaces*, Annals of Mathematics, vol. 36 (1935), pp. 719–724.

‡ Cf. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932, p. 181.

$N=(\sum_{i=1}^{k} a_i^p)^{1/p}$ $(p>1)$; here condition (b) becomes the inequality of Minkowski.*

Suppose now that a finite number of Banach spaces, $B_1, B_2, \cdots, B_k$ are given, and that B is their product. We shall call B a *uniformly convex product* of the B_i if the norm of an element $x=(x^1, x^2, \cdots, x^k)$ of B is defined by

$$\|x\| = N(\|x^1\|, \|x^2\|, \cdots, \|x^k\|),$$

where N is a continuous non-negative function satisfying the conditions (a)-(c).

We now prove

THEOREM 1. *The uniformly convex product of a finite number of uniformly convex Banach spaces is uniformly convex.*†

We see at once that the norm thus defined satisfies the usual rules, including the triangle inequality. Let $B=B_1\times B_2\times\cdots\times B_k$, and let $\{x_i\}=\{(x_i^1, x_i^2, \cdots, x_i^k)\}$, $\{y_i\}=\{(y_i^1, y_i^2, \cdots, y_i^k)\}$ be two sequences of points of B such that $\|x_i\|=\|y_i\|=1$, $\|x_i+y_i\|\to 2$ as $i\to\infty$. We must show that $\|x_i-y_i\|\to 0$.

In the first place we assert that $\lim_{i\to\infty}(\|x_i^j\|-\|y_i^j\|)=0$ $(j=1, 2, \cdots, k)$. Indeed, we have, using conditions (a)-(c),

$$\begin{aligned}\|x_i+y_i\| &= N(\|x_i^1+y_i^1\|, \cdots, \|x_i^k+y_i^k\|)\\ &\leqq N(\|x_i^1\|+\|y_i^1\|, \cdots, \|x_i^k\|+\|y_i^k\|)\\ &\leqq N(\|x_i^1\|, \cdots, \|x_i^k\|)+N(\|y_i^1\|, \cdots, \|y_i^k\|)\\ &= \|x_i\|+\|y_i\| = 2,\end{aligned}$$

so that

$$N(\|x_i^1\|+\|y_i^1\|, \cdots, \|x_i^k\|+\|y_i^k\|)\to 2.$$

Now the last expression is <2 unless $\|x_i^j\|=\|y_i^j\|$ $(j=1, 2, \cdots, k)$ by (b) and our normalization; by the continuity of N our above assertion then follows.

Assume that $\|x_i-y_i\|$ does not $\to 0$. We may then by selecting an appropriate subsequence of i's assume that

---

* See, for example, Hardy, Littlewood and Pólya, *Inequalities*, Cambridge, 1934, p. 31.

† In general, of course, $\|x\|$ will depend upon the x_i themselves, and not merely upon their norms. If this is the case, however, the question of the uniform convexity of the product space is entirely independent of the uniform convexity of the factor spaces, and no theorem of the type of Theorem 1 can be stated.

$$(2)\qquad \begin{cases} \lim_{i\to\infty} \|x_i - y_i\| = \alpha > 0; \\ \lim_{i\to\infty} \|x_i^j\| = \lim_{i\to\infty} \|y_i^j\| = \beta_j \qquad (j = 1, 2, \cdots, k); \\ \lim_{i\to\infty} \|x_i^1 - y_i^1\| = \gamma > 0; \\ \|x_i^1\| > 0, \qquad \|y_i^1\| > 0 \qquad (i = 1, 2, \cdots). \end{cases}$$

Set

$$w_i = \beta_1 x_i^1 / \|x_i^1\|,$$
$$z_i = \beta_1 y_i^1 / \|y_i^1\|.$$

Then

$$\lim_{i\to\infty} \|w_i - x_i^1\| = \lim_{i\to\infty} \|z_i - y_i^1\| = 0,$$

so that

$$\lim_{i\to\infty} \|w_i - z_i\| = \gamma.$$

Then since B_1 is uniformly convex,

$$\limsup_{i\to\infty} \|w_i + z_i\| < 2\beta_1,$$

and hence

$$\limsup_{i\to\infty} \|x_i^1 + y_i^1\| < 2\beta_1.$$

Since by (2)

$$\limsup_{i\to\infty} \|x_i^j + y_i^j\| \leqq 2\beta_j \qquad (j = 2, 3, \cdots, k),$$

it follows that

$$\limsup_{i\to\infty} \|x_i + y_i\| \leqq N\left(\limsup_{i\to\infty} \|x_i^1 + y_i^1\|, \cdots, \limsup_{i\to\infty} \|x_i^k + y_i^k\|\right)$$
$$< N(2\beta_1, \cdots 2\beta_k) = 2N(\beta_1, \cdots, \beta_k) = 2,$$

from which contradiction of our assumption the theorem follows.

3. Spaces L_p and l_p $(p>1)$

We might now attempt to prove the uniform convexity of space l_p by extending the argument of Theorem 1 to an infinite number of factors. We prefer to prove this fact and the corresponding statement for L_p by exhibiting a set of inequalities for these spaces which are in close analogy with the identity (1). We collect these in

THEOREM 2. *For space L_p or l_p, with $p \geqq 2$, the following inequalities between the norms of two arbitrary elements x and y of the space are valid (here q is the conjugate index, $q=p/(p-1)$):*

$$2(\|x\|^p + \|y\|^p) \leqq \|x+y\|^p + \|x-y\|^p \leqq 2^{p-1}(\|x\|^p + \|y\|^p); \tag{3}$$

$$2(\|x\|^p + \|y\|^p)^{q-1} \leqq \|x+y\|^q + \|x-y\|^q; \tag{4}$$

$$\|x+y\|^p + \|x-y\|^p \leqq 2(\|x\|^q + \|y\|^q)^{p-1}. \tag{5}$$

For $1 < p \leqq 2$ these inequalities hold in the reverse sense.

We first point out that for all values of p the right-hand side of (3) is equivalent to the left-hand side, while (4) is equivalent to (5); to see this, set $x+y=\xi$, $x-y=\eta$ and reduce. We make full use of this fact in proving the theorem.

Proof of (4). First consider $1<p\leqq 2$. We commence by showing that for x, y, any two complex numbers, we have

$$|x+y|^q + |x-y|^q \leqq 2(|x|^p + |y|^p)^{q-1}. \tag{6}$$

To prove this, assume that $|x| \geqq |y|$, and divide (6) by $|x|^q$, reducing (6) to*

$$|1+c|^q + |1-c|^q \leqq 2(1+|c|^p)^{q-1}, \tag{7}$$

with $|c| \leqq 1$. Setting $c=\rho e^{i\theta}$, we see by elementary calculus methods that it suffices to consider $\theta=0$; i.e., $0 \leqq c \leqq 1$; (7) is trivial for $c=0$ or $c=1$, so we need only consider $0<c<1$. Making the further transformation $c=(1-z)/(1+z)$ (so that $0<z<1$) we reduce (7) to the form

$$S \equiv \tfrac{1}{2}\{(1+z)^p + (1-z)^p\} - (1+z^q)^{p-1} \geqq 0.$$

Expanding each term of S in its Taylor's series, we have

$$\tfrac{1}{2}\{(1+z)^p + (1-z)^p\} = 1 + \frac{p(p-1)}{2!}z^2 + \frac{p(p-1)(2-p)(3-p)}{4!}z^4 + \cdots + \frac{p(p-1)(2-p)\cdots(2k-1-p)}{(2k)!}z^{2k} + \cdots;$$

$$(1+z^q)^{p-1} = 1 + (p-1)z^q - \frac{(p-1)(2-p)}{2!}z^{2q} + \cdots + \frac{(p-1)(2-p)\cdots(2k-1-p)}{(2k-1)!}z^{(2k-1)q} - \frac{(p-1)(2-p)\cdots(2k-p)}{(2k)!}z^{2kq} + \cdots.$$

* The proof of (7) given here is due to J. S. Frame.

Hence

$$S = \sum_{k=1}^{\infty}\left[\frac{p(p-1)(2-p)\cdots(2k-1-p)}{(2k)!}z^{2k} - \frac{(p-1)(2-p)\cdots(2k-1-p)}{(2k-1)!}z^{(2k-1)q} + \frac{(p-1)(2-p)\cdots(2k-p)}{(2k)!}z^{2kq}\right]$$

$$= \sum_{k=1}^{\infty}\frac{(2-p)(3-p)\cdots(2k-p)}{(2k-1)!}z^{2k}\left[\frac{1-z^{(2k-p)/(p-1)}}{(2k-p)/(p-1)} - \frac{1-z^{2k/(p-1)}}{2k/(p-1)}\right].$$

But $(1-z^t)/t$, for $t>0$ and $0<z<1$, is a decreasing function of t; hence the series for S has non-negative terms, and (7) is established.

Turning now to the proof of (4), we consider first space l_p. Let the two elements considered be $x=(x_1, x_2, \cdots)$, $y=(y_1, y_2, \cdots)$; (4) states that

$$\left[\sum_{i=1}^{\infty}|x_i+y_i|^p\right]^{q/p} + \left[\sum_{i=1}^{\infty}|x_i-y_i|^p\right]^{q/p} \leqq 2\left[\sum_{i=1}^{\infty}(|x_i|^p+|y_i|^p)\right]^{q-1}. \tag{8}$$

Now one form of Minkowski's Inequality states that if A_i, B_i are any two sets of non-negative numbers, finite or infinite in number, and $0<s\leqq 1$, then

$$(\sum A_i^s)^{1/s} + (\sum B_i^s)^{1/s} \leqq (\sum (A_i+B_i)^s)^{1/s},$$

proper account being taken of convergence. Setting

$$\frac{p}{q} = s, \quad |x_i+y_i|^q = A_i, \quad |x_i-y_i|^q = B_i,$$

we infer that the left side of (8) is

$$\leqq \left[\sum_{i=1}^{\infty}(|x_i+y_i|^q+|x_i-y_i|^q)^{p/q}\right]^{q/p}$$

which by (6) is

$$\leqq \left[\sum_{i=1}^{\infty}(2\{|x_i|^p+|y_i|^p\}^{q/p})^{p/q}\right]^{q/p} = \left[\sum_{i=1}^{\infty}2^{p/q}\{|x_i|^p+|y_i|^p\}\right]^{q/p}.$$

Since $q/p=q-1$, this is our result; (4), then, stands proved for l_p $(1<p\leqq 2)$.

To extend this result to space L_p $(1<p\leqq 2)$, let $[0, 1]$ be the interval over which the functions of our space are to be defined. We consider first two func-

tions $x(t)$, $y(t)$ which are step functions on a division of $[0, 1]$ into equal parts. It is readily verified that for such functions the relation (4) reduces to the l_p case already treated, and as these functions form a dense set in L_p the result follows by continuity of the norm.

Consider now (4) for $p>2$. Again let x, y be any two elements of l_p; the relation which we must prove is (8) with the sense reversed. Letting A_i, B_i, and s have the same values as above, and again applying Minkowski's Inequality, which is now reversed in sense since s exceeds 1, we conclude that the left side of (8) is

$$\geqq \left[\sum_{i=1}^{\infty} (|x_i + y_i|^q + |x_i - y_i|^q)^{p/q} \right]^{q/p}. \tag{9}$$

Now (5) (or its equivalent, (4)) has already been proved for $p<2$; hence employing it for complex numbers we have that for $p>2$,

$$|x_i + y_i|^q + |x_i - y_i|^q \geqq 2(|x_i|^p + |y_i|^p)^{q-1}.$$

From this we have that (9) is

$$\geqq \left[\sum_{i=1}^{\infty} (2\{|x_i|^p + |y_i|^p\}^{q-1})^{p/q} \right]^{q/p} = 2\left[\sum_{i=1}^{\infty} (|x_i|^p + |y_i|^p) \right]^{q-1},$$

our result for l_p. The passage to space L_p is carried out exactly as in the case for $1<p\leqq 2$, with which the demonstration of relations (4) and (5) is complete.

Proof of (3). We take $p\geqq 2$, and consider the right-hand inequality. We show that this is implied by (5): that for $a, b\geqq 0$ we have

$$2(a^q + b^q)^{p-1} \leqq 2^{p-1}(a^p + b^p). \tag{10}$$

For suppose $a\leqq b>0$, which entails no real loss of generality; dividing (10) by $b^{q(p-1)} = b^p$ we obtain

$$2(c^q + 1)^{p-1} \leqq 2^{p-1}(c^p + 1) \qquad (0 \leqq c \leqq 1),$$

or

$$2^{p-2} \frac{c^p + 1}{(c^q + 1)^{p-1}} \geqq 1,$$

which, being raised to the power $1/p$, gives

$$H(c) \equiv 2^{(p-2)/p} \frac{(c^p + 1)^{1/p}}{(c^q + 1)^{1/q}} \geqq 1.$$

$H(1) = 1$, and the result then follows by noting that dH/dc is <0 for the in-

terval considered. For $1<p\leqq 2$ the inequality (10) holds in the reverse sense, the proof being identical. This completes the proof of Theorem 2.

COROLLARY. *For $p>1$ the spaces L_p and l_p are uniformly convex.*

It is of some interest to compute the function $\delta(\epsilon)$ of Definition 1 explicitly in this case. Consider $p\geqq 2$, and set $\|x\|=\|y\|=1$ in (3);

$$\|x+y\|^p+\|x-y\|^p\leqq 2^p.$$

Then if $\|x-y\|\geqq\epsilon$ $(0<\epsilon\leqq 2)$, we see that

$$\left\|\frac{x+y}{2}\right\|\leqq\left[1-\left(\frac{\epsilon}{2}\right)^p\right]^{1/p},$$

so that an admissible value for $\delta(\epsilon)$ is $1-[1-(\epsilon/2)^p]^{1/p}$. For $1<p\leqq 2$, we have $\delta(\epsilon)=1-[1-(\epsilon/2)^q]^{1/q}$, by (4).

4. A "STRONG" TRIANGLE RELATION

For uniformly convex spaces the ordinary "triangular inequality," which states that if $y=\sum_{i=1}^k x_i$ then $\|y\|\leqq\sum_{i=1}^k\|x_i\|$, can be replaced by a considerably stronger inequality.

For two non-vanishing vectors, x and y, of the space B, we define a generalized "angle" between the vectors, which we denote by $\alpha[x, y]$, as follows:

$$\alpha[x, y]=\left\|\frac{x}{\|x\|}-\frac{y}{\|y\|}\right\|.$$

We note the following properties of α:

(i) $\alpha[x, y]=\alpha[y, x]$; $0\leqq\alpha[x, y]\leqq 2$.

(ii) $\alpha[x, y]\leqq\alpha[x, z]+\alpha[z, y]$.

(iii) $\alpha[x, y]=0$ is equivalent to the condition $y=cx$ $(c>0)$.

We now state our inequality in

THEOREM 3. *Let $x_1, x_2, \cdots, x_k, y$ be non-vanishing elements of B, a uniformly convex Banach space, with*

$$y=\sum_{i=1}^k x_i.$$

Let $\alpha_i=\alpha[x_i, y]$ $(i=1, 2, \cdots, k)$. Then

$$\|y\|\leqq\sum_{i=1}^k[1-2\delta(\alpha_i)]\|x_i\|, \tag{11}$$

where δ is the function of Definition 1, and $\delta(0)=0$.

Since for each i we have

$$\left\| \frac{x_i}{\|x_i\|} - \frac{y}{\|y\|} \right\| = \alpha_i,$$

then by Definition 1, for each i,

$$\left\| \frac{x_i}{\|x_i\|} + \frac{y}{\|y\|} \right\| \leqq 2 - 2\delta(\alpha_i),$$

or

$$\left\| \|y\| x_i + \|x_i\| y \right\| \leqq (2 - 2\delta(\alpha_i))\|x_i\| \, \|y\|;$$

summing these inequalities over i, we obtain

$$\sum_{i=1}^{k} \left\| \|y\| x_i + \|x_i\| y \right\| \leqq \sum_{i=1}^{k} (2 - 2\delta(\alpha_i))\|x_i\| \, \|y\|.$$

If we replace the sum of norms on the left by the norm of the sum, this becomes

$$\left\| \|y\| y + y \sum_{i=1}^{k} \|x_i\| \right\| \leqq \|y\| \sum_{i=1}^{k} (2 - 2\delta(\alpha_i))\|x_i\|,$$

or, since the left side equals $\|y\| \, (\|y\| + \sum_{i=1}^{k} \|x_i\|)$,

$$\|y\| + \sum_{i=1}^{k} \|x_i\| \leqq 2 \sum_{i=1}^{k} \|x_i\| - 2 \sum_{i=1}^{k} \delta(\alpha_i)\|x_i\|.$$

Rearranging the terms we have the inequality of the theorem.

It will be noticed that in a uniformly convex space, if the sum of the lengths of two sides of a triangle is equal to the length of the third side, the triangle is degenerate, and conversely;* that is, if x and y are $\neq 0$, and $\|x+y\| = \|x\| + \|y\|$, then $x = cy (c > 0)$: this statement is a consequence of (11) and (iii). We can show from this that in Definition 1 the value of $\delta(2)$ may be assumed to be 1. Indeed, suppose that $\|x\| = \|y\| = 1$, and $\|x-y\| = 2$. Then

$$\|x + (-y)\| = \|x\| + \| - y\|;$$

hence by the property just noted, $x = -y$, and the value $\delta(2) = 1$ is admissible. We can now state as a further property of the function $\alpha[x, y]$ that

(iv) $\alpha[x, y] = 2$ is equivalent to the condition $y = cx (c < 0)$.

Using the particular value of $\delta(\epsilon)$ obtained above for $L_p[l_p]$ we have the following strengthening of Minkowski's Inequality:

* We say that a space in which this condition is satisfied is "strictly convex."

COROLLARY. *Let $x_1, x_2, \cdots, x_k, y$ be non-vanishing elements of L_p or l_p $(p \geqq 2)$, $y = \sum_{i=1}^{k} x_i$. If $\alpha_i = \alpha[x_i, y]$, then*

$$\|y\| \leqq \sum_{i=1}^{k} [2(1 - (\alpha_i/2)^p)^{1/p} - 1]\|x_i\|. \tag{12}$$

For $1 < p \leqq 2$, p must be replaced by q in this inequality.

We point out that in inequalities (11) and (12) the factors associated with the $\|x_i\|$ may assume any values between -1 and 1 inclusive. These relations are obviously reminiscent of the corresponding relation between vectors in the Euclidean plane, $\|y\| = \sum \cos a_i \|x_i\|$, a_i = angle between the vectors x_i and y.

For use in a later argument we state the following corollary to Theorem 3.

COROLLARY. *Let $y, x_1, x_2, \cdots, x_k, w_1, \cdots, w_l$ be non-vanishing elements of a uniformly convex space, with $y = \sum_{i=1}^{k} x_i + \sum_{i=1}^{l} w_i$, and with $\alpha[x_i, y] > \epsilon > 0$ $(i = 1, 2, \cdots, k)$. Then*

$$\sum_{i=1}^{k} \|x_i\| \leqq \frac{1}{2\delta(\epsilon)} \left(\sum_{i=1}^{k} \|x_i\| + \sum_{i=1}^{l} \|w_i\| - \|y\| \right).$$

This follows immediately from

$$\|y\| \leqq \sum_{i=1}^{k} (1 - 2\delta(\epsilon))\|x_i\| + \sum_{i=1}^{l} \|w_i\|;$$

which itself is implied by Theorem 3.

5. FUNCTIONS FROM A EUCLIDEAN TO A UNIFORMLY CONVEX SPACE

A function of bounded variation from a Euclidean space into a uniformly convex Banach space is differentiable at almost all points; we shall formulate and prove this statement in precise fashion as our next theorem. As we have noted above, Bochner has shown that this theorem is not true for Banach spaces in general. Bochner's example is a function from one-dimensional Euclidean space into M, the space of bounded functions. The following example, in which the range space is taken to be space L, the space of all summable functions $f(s)$ on the interval $0 \leqq s \leqq 1$, shows that this phenomenon may still occur in separable spaces.

To each point t, $0 \leqq t \leqq 1$, let correspond the element $\phi_t \epsilon L$ defined as follows:

$$\phi_t = \phi_t(s) = \begin{cases} 1 & (s \leqq t), \\ 0 & (s > t). \end{cases}$$

The reader will easily verify that the norm in L of the difference quotient

$$\Delta(h, t) = \frac{\phi_{t+h} - \phi_t}{h} \qquad (h \neq 0)$$

is identically equal to unity, and hence that the point function ϕ_t, since it satisfies a Lipschitz condition, is of bounded variation and indeed is absolutely continuous. However, as $h \to 0$, $\Delta(h, t)$ approaches no element of L even in the weak sense,* and hence ϕ_t has no (strong) derivative at any point; that is, $\lim_{h \to 0} \Delta(h, t)$ does not exist.

In our proof of the abstract form of the Lebesgue theorem we employ the terminology and definitions of Saks.† The basic method of the proof, however, is contained in a proof of Tonelli‡ that rectifiable plane curves admit tangents almost everywhere.

We recall a few definitions.

By an *interval* of the n-dimensional Euclidean space $\mathcal{E}_n$ is meant the set of points $(x_1, x_2, \cdots, x_n)$ satisfying the inequalities $a_i \leqq x_i \leqq b_i$, where (a_i, b_i) are n number pairs with $a_i < b_i$. In case all $b_i - a_i$ are equal, the interval is called a *cube*, and will always be denoted by I. By an *elementary figure*, R, is meant the sum of a finite number of intervals (or the null set). In order that the set of elementary figures may be closed under our manipulations, the notions of difference and intersection of two figures, denoted symbolically by $\ominus$ and $\odot$, are defined as

$$R_1 \ominus R_2 = \overline{(R_1 - R_2)^0},$$
$$R_1 \odot R_2 = \overline{(R_1 \cdot R_2)^0},$$

where E^0, $\overline{E}$ denote the interior and closure respectively of the set E. We denote the measure of the figure by $|R|$.

A function $F(R)$ which makes correspond to each elementary figure, contained in a fixed figure R_0, an element in a linear space, is called *additive* if

$$F(R_1 + R_2) = F(R_1) + F(R_2)$$

whenever R_1 and R_2 are non-overlapping; that is, the set $R_1 \cdot R_2$ is of measure zero. $F(R)$ is said to be of *bounded variation* in R_0 if for any finite set of non-overlapping figures in R_0: $R_1, R_2, \cdots, R_k$, the sum $\sum_{i=1}^{k} \|F(R_i)\|$ is less than a fixed bound. $F(R)$ is *absolutely continuous*§ in R_0 if given $\epsilon > 0$ there corresponds a $\delta_\epsilon > 0$ such that if $|R| < \delta_\epsilon$, then $\|F(R)\| < \epsilon$.

* See Banach, loc. cit., p. 136.

† S. Saks, *Théorie de l'Intégrale*, Chapters I–III, esp. p. 47.

‡ Tonelli, *Fondamenti di Calcolo della Variazione*, pp. 48–56.

§ These last two definitions, because of the setting, differ slightly in form from those of Saks, but are equivalent to his in case the range space is the set of real numbers.

Finally, $F(R)$ is said to be *differentiable* at a point p if the limit

$$F'(p) = \lim_{|I|\to 0} \frac{F(I)}{|I|}$$

exists, where I is an arbitrary cube containing p.

THEOREM 4. *Let $F(R)$ be an additive function of elementary figures, defined for the figures within a fixed figure R_0, and assuming values in a uniformly convex Banach space B. Let F be of bounded variation in R_0. Then F is differentiable almost everywhere in R_0.*

We first note that in proving the existence of $F'(p)$ we may without loss of generality consider only those functions $F(R)$ such that

$$\|F(R)\| \geqq |R| \tag{13}$$

for all R. Indeed, we have to prove that

$$\lim_{|I|\to 0} \frac{F(I)}{|I|} \tag{14}$$

exists, where I is an arbitrary cube containing the point p. Let $B_1 = B \times \mathcal{E}_1$, where $\|(x, t)\| = [\|x\|^2 + t^2]^{1/2}$, and consider the function of figures $F_1(R) = (F(R), |R|)$, whose range is in B_1. The function F_1 satisfies our conditions, since B_1 is uniformly convex by Theorem 1, and also (13); hence, assuming the theorem proved in this case, we have almost everywhere that

$$F_1'(p) = \lim_{|I|\to 0} \frac{(F(I), |I|)}{|I|} \qquad (p \epsilon I)$$

exists. This, however, implies the existence of (14), which justifies our preliminary remark.

We first show that the limit

$$\lim_{|I|\to 0} \frac{F(I)}{\|F(I)\|} \qquad (p \epsilon I) \tag{15}$$

exists for almost all p.

Given a point p, and $\epsilon > 0$, we define

$$A(p, \epsilon) = \sup \alpha[F(I), F(I')],$$

where I, I' are arbitrary cubes with measure less than ϵ containing p. Let $A(p) = \lim_{\epsilon\to 0} A(p, \epsilon)$. A necessary and sufficient condition that (15) exist is that $A(p) = 0$.

Let ϵ, $\epsilon'>0$ be fixed. Let $R_0=\sum_{i=1}^{k}R_i$, R_i non-overlapping elementary figures. By assumption the sums $\sum_{i=1}^{k}\|F(R_i)\|$ have a least upper bound M for all choices of the R_i; we fix this choice so that

$$M-\sum_{i=1}^{k}\|F(R_i)\|<\epsilon'. \tag{16}$$

Let $E=\underset{\rho}{E}\{A(p)>2\epsilon\}$; then $m^*E=\sum_{i=1}^{k}m^*ER_i^{\rho}$. Each set ER_i^{ρ} is covered (in the sense of Vitali)† by a set of cubes I^i, $\subset R_i^{\rho}$, with

$$\alpha[F(I^i),F(R_i)]>\epsilon. \tag{17}$$

Then by Vitali's covering theorem there is a finite set of these, $\{I_j^i\}$ $(j=1,2,\cdots,n_i)$, disjunct, with

$$\sum_{j=1}^{n_i}|I_j^i|>\tfrac{1}{2}m^*ER_i^0=\tfrac{1}{2}m^*ER_i \qquad (i=1,2,\cdots,k). \tag{18}$$

Now by additivity

$$F(R_i)=\sum_j F(I_j^i)+F\Big(R_i\ominus\sum_j I_j^i\Big),$$

so that by (17) and the second corollary to Theorem 3 we infer

$$\sum_j\|F(I_j^i)\|\leqq\frac{1}{2\delta(\epsilon)}\Big[\sum_j\|F(I_j^i)\|+\Big\|F\Big(R_i\ominus\sum_j I_j^i\Big)\Big\|-\|F(R_i)\|\Big],$$

or, adding these inequalities over i,

$$\sum_i\sum_j\|F(I_j^i)\|$$
$$\leqq\frac{1}{2\delta(\epsilon)}\Big[\sum_i\sum_j\|F(I_j^i)\|+\sum_i\Big\|F\Big(R_i\ominus\sum_j I_j^i\Big)\Big\|-\sum_i\|F(R_i)\|\Big].$$

The sum of the first two terms in the bracket is $\leqq M$; hence by (13) and (16) we can conclude that

$$\sum_i\sum_j|I_j^i|\leqq\frac{\epsilon'}{2\delta(\epsilon)},$$

which in conjunction with (18) gives us

$$m^*E=\sum_i m^*ER_i<2\sum_i\sum_j|I_j^i|\leqq\frac{\epsilon'}{\delta(\epsilon)}.$$

† For a definition of a Vitali covering, and a proof of Vitali's theorem, see, for example, Saks, loc. cit., pp. 33–36.

Since E is independent of ϵ', this implies $mE=0$, which in turn implies $A(p)=0$ almost everywhere. The existence of the limit (15) almost everywhere has now been established.

Now to arrive at the conclusion that the limit (14) exists, we again form the function $F_1(R)=(F(R), |R|)$ with range in space $B_1=B\times\mathcal{E}_1$. Since (15) holds almost everywhere, we have the existence of

$$\lim_{|I|\to 0}\frac{F_1(I)}{\|F_1(I)\|}=\lim_{|I|\to 0}\frac{(F(I), |I|)}{(\|F(I)\|^2+|I|^2)^{1/2}},$$

and hence that of both limits

$$\lim_{|I|\to 0}\frac{F(I)}{(\|F(I)\|^2+|I|^2)^{1/2}}, \qquad \lim_{|I|\to 0}\frac{|I|}{(\|F(I)\|^2+|I|^2)^{1/2}}$$

except for a null set. If the second limit does not vanish except for a null set, then (14) will follow. Now where the second limit vanishes, we have

$$\lim_{|I|\to 0}\frac{\|F(I)\|}{|I|}=+\infty. \tag{19}$$

But given any $N>0$ let E_N be the set on which

$$\limsup_{|I|\to 0}\frac{\|F(I)\|}{|I|}>N.$$

Then corresponding to every point $p\epsilon E_N$ there is a set of cubes $I^i\ni p$ with $|I^i|\to 0$ and $\|F(I^i)\|>N|I^i|$; by another application of Vitali's theorem there is a finite set of these, non-overlapping, with

$$\sum|I_j|\geqq\tfrac{1}{2}m^*E_N,$$

which yields the estimate

$$m^*E_N\leqq 2\sum|I_j|<\frac{2}{N}\sum\|F(I_j)\|\leqq\frac{2M}{N}.$$

Thus (19) can only occur on a null set, and (14) exists almost everywhere. This completes the proof of Theorem 4.

THEOREM 5. *Under the hypotheses of Theorem* 4, *the derivative of* $F(R)$ *is integrable in the sense of Bochner.*

Let I_0 be a cube containing R_0, and for each positive integer k let I_0 be subdivided in symmetrical fashion into 2^{nk} sub-cubes, I_j^k. Now in I_0 define the functions $F_k(p)$ as follows: if p is interior to a cube $I_j^k\subset R_0$, define $F_k(p)=F(I_j^k)/|I_j^k|$; for all other points of I_0 let $F_k(p)$ vanish. Then almost

everywhere in R_0 we have $\lim_{k\to\infty} F_k(p) = F'(p)$, and, since the F_k are "finite valued functions," $F'(p)$ is measurable as a consequence. Now to show that it is integrable we need only demonstrate* that the real function $\|F'(p)\|$, which is also measurable, is integrable. Since $\|F_k(p)\| \to \|F'(p)\|$ almost everywhere, this will follow from the Fatou lemma† if the integrals $\int_{R_0} \|F_k(p)\|$ are bounded. But for each k,

$$\int_{R_0} \|F_k(p)\| \leqq \int_{I_0} \|F_k(p)\| = \sum_{I_j^k \subset R_0} \frac{\|F(I_j^k)\|}{|I_j^k|} |I_j^k| \leqq M,$$

which completes the proof.

THEOREM 6. *Let the hypotheses of Theorem* 4 *be satisfied, and in addition let* $F(R)$ *be absolutely continuous in* R_0. *Then* F *is the integral of its derivative; for every elementary figure* $R \subset R_0$,

$$F(R) = \int_R F'(p).$$

That $F'(p)$ exists almost everywhere, and is a summable function, we have already seen; set $G(R) = \int_R F'(p)$ for each figure $R \subset R_0$. Bochner has shown‡ that $G'(p)$ exists and equals $F'(p)$ almost everywhere; moreover G is additive and absolutely continuous. Then the function $H = F - G$ is an additive absolutely continuous function of figures whose derivative vanishes almost everywhere. We assert that from this it follows that H vanishes identically.§ Indeed, let R be any elementary figure, and let Ξ be the set on which $H'(p) = 0$. Let $\epsilon > 0$ be chosen, and let $\delta(\epsilon)$ be the function whose existence is asserted in the definition of absolute continuity. The set $R \cdot \Xi$ is covered (Vitali) by a set of cubes I for each of which we have

$$H(I) < \delta(\epsilon) |I|. \tag{20}$$

Then by Vitali's theorem there is a finite set of these cubes, $\{I_j\}$, disjunct, such that

$$mR \cdot \Xi - \frac{\delta(\epsilon)}{2} < m(R \cdot \Xi \cdot \sum I_j) \leqq m \sum I_j < mR \cdot \Xi + \frac{\delta(\epsilon)}{2},$$

---

* Since Bochner's integral is absolutely convergent.

† Cf. Schlesinger and Plessner, *Lebesguesche Integrale und Fouriersche Reihen*, Berlin, 1926, p. 91.

‡ See the first cited paper of Bochner, p. 269.

§ As the proof shows, this assertion is true if B is any Banach space. In fact, since Theorem 5 employs the assumption of uniform convexity only to assert the existence of $F'(R)$, we can say that if $F(R)$ is an additive absolutely continuous function from $\mathfrak{E}_n$ into any Banach space, which is differentiable almost everywhere, then $F'(p)$ is integrable (Bochner), and F is its indefinite integral.

and as Ξ includes almost all points of R_0, we may rewrite this suppressing the factor Ξ, as

$$mR - \frac{\delta(\epsilon)}{2} < mR \odot \sum I_i \leqq m \sum I_i < mR + \frac{\delta(\epsilon)}{2}. \tag{21}$$

Now

$$(\sum I_i \ominus [R \odot \sum I_i]) + R = \sum I_i + (R \ominus [R \odot \sum I_i]),$$

so that by additivity and the triangle relation one has

$$\begin{aligned} \|H(R)\| \leqq \sum \|H(I_i)\| + \|H(R \ominus [R \odot \sum I_i])\| \\ + \|H(\sum I_i \ominus [R \odot \sum I_i])\|. \end{aligned} \tag{22}$$

From (21) we deduce the inequalities

$$\begin{cases} m(R \ominus [R \odot \sum I_i]) = mR - mR \odot \sum I_i < \delta(\epsilon), \\ m(\sum I_i \ominus [R \odot \sum I_i]) = m \sum I_i - mR \odot \sum I_i < \delta(\epsilon); \end{cases} \tag{23}$$

combining the relations (20), (21), (22), and (23), and remembering that $mR' < \delta(\epsilon)$ implies $\|H(R')\| < \epsilon$, we infer that

$$\|H(R)\| < \delta(\epsilon)\left[mR + \frac{\delta(\epsilon)}{2}\right] + 2\epsilon.$$

As we may assume that $\delta(\epsilon) \leqq \epsilon$, this is arbitrarily small with ϵ, $\|H(R)\|$ must vanish, and the proof is complete.

COROLLARY. *An additive function of elementary figures, with values in a uniformly convex space, which is absolutely continuous and of bounded variation, can be extended to all measurable sets while preserving these properties, and in only one way.*

By a standard argument we can now prove, for point functions in the one-dimensional case,

THEOREM 7. *A function $\phi(t)$ from an interval $a \leqq t \leqq b$ into B, a uniformly convex space, which is of bounded variation, is (strongly) differentiable almost everywhere. The derivative, $\phi'(t)$, is integrable (Bochner). If in addition $\phi(t)$ is absolutely continuous, then**

$$\phi(t) = \phi(a) + \int_a^t \phi'(t)dt.$$

We omit the proof, which presents no difficulty.

* Compare G. Birkhoff, loc. cit., Theorem 26, which is the special case of Theorem 7 in which B is Hilbert space. The last assertion of this theorem appears to be in error.

The notion of a "continuous rectifiable curve" in a uniformly convex space may be defined in a manner quite analogous to the procedure for ordinary plane curves, and by using the "angle" α defined above, the notion of a "tangent" to such a curve may be made precise. We omit the details, but state without proof the following theorem, the demonstration of which may be made to depend directly on the preceding theory:

THEOREM 8. *A continuous rectifiable curve in a uniformly convex space possesses a tangent at almost all of its points.*

6. A GENERALIZATION. CONCLUSION

It has been called to the attention of the author by J. von Neumann that the "differentiability properties" of the space expressed in the last five theorems remain true under a somewhat weaker assumption. It suffices to assume, not that the space B is uniformly convex, but merely that there can be defined a norm in the space, equivalent to the given norm (i.e., such that the notion of limit is the same with both), with respect to which the space is uniformly convex. The only point that is not immediately obvious in the verification of this statement is that the definitions of bounded variation and absolute continuity are equivalent with the different norms, and this is clear if we notice that when $\|x\|_1$, $\|x\|_2$ are two equivalent norms, the ratio $\|x\|_1 / \|x\|_2$ is bounded for $\|x\|_2 \neq 0$.

It follows from this that these properties are possessed by all finite dimensional Euclidean spaces with an arbitrary "Minkowski norm"; that is, a norm determined by any convex region centrally symmetric about the origin; since any such metric is clearly equivalent to the Euclidean metric. For example, all spaces $l_1^{(k)}$ with a finite number, k, of coordinates, possess these properties.

It may be proved directly without a great deal of difficulty that the space l_1 also possesses these properties. The proof rests essentially on the denumerable coordinate system in space l_1. If a function $\phi(t)$, say from a real segment into (real) l_1, is of bounded variation, then each of its components $x_i(t)$ is a real function of bounded variation, and the sum of their variations is finite. Then $\phi(t)$ may be decomposed into the difference of two "monotone" functions by decomposing each of the $x_i(t)$; by means of a theorem of Fubini, that a convergent series of monotonically increasing functions is term-wise differentiable, the result follows.

A question which will naturally arise in the reader's mind is whether "differentiability theorems" of the above nature can be proved under the still weaker assumption that the space B in question is strictly convex,* or

* See the definition above.

that a norm equivalent to the given norm exists with respect to which B is strictly convex. The answer to this question is negative: indeed, the essential role played by the uniformity assumption is clearly exhibited by the following theorem.

THEOREM 9. *Any separable Banach space may be given a new norm, equivalent to the original norm, with respect to which the space is strictly convex.*

We first demonstrate that the theorem is true of space C (the space of continuous functions $f(t)$ in the interval $0 \leqq t \leqq 1$, where $\|f\| = \max |f(t)|$). Let $\{t_n\}$ be a dense sequence of points in $(0, 1)$; then as F. J. Murray has remarked, the sequence of bounded linear functionals $F_n(f) = f(t_n)$ forms a "total" set in the space $\overline{C}$ conjugate to C; that is, the only element of space C for which all of these functionals vanish is the zero element. Now let the space C be renormed with the following definition:

$$\|f\|_1 = \left[\|f\|^2 + \sum_{n=1}^{\infty} \frac{1}{2^{2n}} |F_n(f)|^2\right]^{1/2};$$

it is readily verified that $\|f\|_1$ is a norm equivalent to $\|f\|$. But suppose that for two non-vanishing elements f and g we have $\|f+g\|_1 = \|f\|_1 + \|g\|_1$. This implies

$$\left[(\|f\| + \|g\|)^2 + \sum_{n=1}^{\infty} \frac{1}{2^{2n}} |F_n(f) + F_n(g)|^2\right]^{1/2}$$
$$= \left[\|f\|^2 + \sum_{n=1}^{\infty} \frac{1}{2^{2n}} |F_n(f)|^2\right]^{1/2} + \left[\|g\|^2 + \sum_{n=1}^{\infty} \frac{1}{2^{2n}} |F_n(g)|^2\right]^{1/2},$$

and hence there exists a constant $c > 0$ with $cF_n(f) = F_n(g)$ $(n = 1, 2, \cdots)$. As the sequence $\{F_n\}$ is total, this implies that $cf = g$; i.e., the renormed space is strictly convex.

Now let B be any separable Banach space. Banach has shown* that there is a closed linear manifold of space C which is in one-to-one linear isometric correspondence with B. Then the above defined renorming of C furnishes a means of renorming the space B in the manner required, which finishes the demonstration of the theorem.

In the opposite direction we have

THEOREM 10. *None of the following spaces can be renormed so as to be uniformly convex:*† *space L; space M (bounded functions); space C (continuous functions); space m (bounded sequences); space c (convergent sequences).*

* Banach, loc. cit., p. 185.

† For the precise definition of these spaces see, for example, Banach, loc. cit., pp. 9–12.

It need only be verified that Theorems 4–8 do not hold in the spaces enumerated. Our above example, and the example given by Bochner, dispose of spaces L and M, and the property of space C used in the last theorem, that it is a "universal space" for all separable Banach spaces, together with our example in space L, furnishes an immediate proof for this space. The following example, of an absolutely continuous (even Lipschitzean) function, from the interval $0 \leqq t \leqq 1$ into space c, which is not differentiable, even in the weak sense, at any point, settles the matter for space c; as this space forms a closed linear manifold of space m, the same example will do for the latter space.

To construct the example, we first form the sequence $f_n(t)$ of real functions on (0, 1) as follows:

$$f_1(t) = \begin{cases} t & (0 \leqq t \leqq \frac{1}{2}), \\ 1 - t & (\frac{1}{2} \leqq t \leqq 1); \end{cases}$$

$$f_n(t) = \begin{cases} \frac{1}{2} f_{n-1}(2t) & (0 \leqq t \leqq \frac{1}{2}), \\ f_n(t - \frac{1}{2}) & (\frac{1}{2} \leqq t \leqq 1). \end{cases}$$

The function $\phi(t)$ from (0, 1) into space c is now defined by setting $\phi(t) = \{f_1(t), f_2(t), \cdots\}$. We leave to the reader the simple verification that $\phi(t)$ possesses the required properties.

The conjugate space of c is the separable space l_1. We see, then, that this property, separability of the conjugate space, is neither necessary nor sufficient for differentiability properties of the type discussed above, thus answering in the negative a question posed by G. Birkhoff* in his paper referred to.

* Loc. cit. p. 378.

INSTITUTE FOR ADVANCED STUDY AND PRINCETON UNIVERSITY,
PRINCETON, N. J.

REMARKS ON THE PRECEDING PAPER OF JAMES A. CLARKSON*

BY

NELSON DUNFORD AND ANTHONY P. MORSE

1. In the preceding paper Clarkson has introduced the interesting concept of a Banach space X with a uniformly convex norm and has shown that for such spaces the following theorem holds.

THEOREM. *If the additive function $F(R)$ which is defined for elementary figures R contained in a fixed figure R_0 in Euclidean space of n dimensions and has its values in the Banach space X, is of bounded variation,† then it has a derivative $F'(P)$ for almost all points P in R_0. $F'(P)$ is summable‡ on R_0 and if $F(R)$ is absolutely continuous§ then for every elementary figure R in R_0 we have*

$$F(R) = \int_R F'(P)dP.$$

In this paper it is shown that the theorem holds for all Banach spaces X with a base¶ $\{\phi_i\}$ which satisfies the following postulate:

(A) *If $a_1, a_2, \cdots$ is any sequence of real numbers such that* $\sup_n \left\| \sum_{i=1}^n a_i\phi_i \right\| < \infty$, *then the series $\sum_{i=1}^\infty a_i\phi_i$ converges.*

It is obvious that l_p $(p \geqq 1)$ or any Hilbert space satisfies (A).

In §3 it is shown that L_p $(p>1)$ does likewise. The method of proof is entirely different from that of Clarkson. In §4 it is shown that if X is any Banach space having the property that every function on a linear interval to X, which satisfies a Lipschitz condition, is differentiable almost everywhere, then also every function of bounded variation from the linear interval to X is differentiable almost everywhere and its derivative is summable.

2. **Proof of the theorem.** It should first be noted that it is no loss of generality to assume besides (A) the property

(B) $$\left\| \sum_{i=1}^n a_i\phi_i \right\| \leqq \left\| \sum_{i=1}^{n+1} a_i\phi_i \right\| \quad \textit{for any constants } a_i,\ i = 1, 2, \cdots.$$

* Presented to the Society, April 11, 1936; received by the editors February 16, 1936.

† I.e., $\sum_{i=1}^k \|F(R_i)\|$ is bounded for all finite sets of non-overlapping elementary figures $R_1, \cdots, R_k$ contained in R_0.

‡ Here, as well as elsewhere in this note, the concept of summability is that of Bochner or of Dunford. The two notions are equivalent.

§ I.e., $\lim_{|R|=0} F(R) = 0$.

¶ See Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932, p. 110.

This is made evident by the following two considerations. If a Banach space Y is isomorphic* to X then the theorem holds for X if and only if it holds for Y. Now associated with X is a space Y composed of all sequences $y=\{\eta_i\}$ such that $\sum_{i=1}^{\infty}\eta_i\phi_i$ converges. If the metric in Y is defined by $\|y\|=\sup_n\|\sum_{i=1}^{n}\eta_i\phi_i\|$ then† Y is a Banach space and isomorphic to X. By taking $y_k=\{\eta_i^k\}$ where $\eta_i^k=0$ if $k\neq i$, $\eta_i^i=1$, it is easily seen that y_i forms a base for Y which has the properties (A) and (B). Thus it will be assumed in what follows that X has the properties (A) and (B). Now

$$F(R)=\sum_{i=1}^{\infty}a_i(R)\phi_i,$$

where the coefficients $a_i(R)$ are given by means of the limited linear functionals‡ T_i defined on X according to the equation $a_i(R)=T_iF(R)$. Thus it is immediate that the functions $a_i(R)$ are additive, real functions of bounded variation with summable derivatives $a_i'(P)$. The functions $F_n(R)=\sum_{i=1}^{n}a_i(R)\phi_i$ thus have derivatives $F_n'(P)=\sum_{i=1}^{n}a_i'(P)\phi_i$ summable on R_0. Let $V(F_n, R)$ be the total variation of F_n on R, then the positive real function $S_n(R)=V(F_n, R)$ is additive and of bounded variation on R_0 with $S_n(R)\geqq\|F_n(R)\|$. Hence $S_n'(P)\geqq\|F_n'(P)\|$ and thus

$$(1)\qquad V(F_n, R)=S_n(R)\geqq\int_R S_n'(P)dP\geqq\int_R\|F_n'(P)\|dP.$$

Now from postulate (B) and (1)

$$(2)\qquad \|F_n'(P)\|\leqq\|F_{n+1}'(P)\|,\qquad \int_{R_0}\|F_n'(P)\|dP\leqq V(F_n, R_0)\leqq V(F, R_0).$$

If we let $b(P)=\lim_n\|F_n'(P)\|$, the inequality (2) shows that $b(P)$ is summable and hence finite almost everywhere. Postulate (A) then insures the convergence of the series

$$G(P)=\sum_{n=1}^{\infty}a_n'(P)\phi_n$$

for almost all P in R_0. Since $\|G(P)\|=b(P)$ and $G(P)$ is measurable it is summable on R_0. It will now be shown that $G(P)$ is the derivative of $F(R)$. To do this we need the following lemmas which, in the case of real-valued functions, are well known.

---

* Banach, loc. cit., p. 180.

† Banach, loc. cit., p. 111.

‡ Banach, loc. cit., p. 111.

LEMMA 1. *Let $G(R)$ be an additive function of bounded variation defined for elementary figures R contained in a fixed figure R_0 and with values in an arbitrary Banach space. Then $G(R)$ has a derivative equal to zero at almost all points of R_0 if and only if for every $\epsilon>0$ there is an open set E in R_0^0 with measure less than ϵ such that $V(G, E)=V(G, R_0^0)$.*

By the variation $V(G, D)$ of G on an open set D in R_0 is meant the upper bound of all finite sums $\sum_{i=1}^{k}\|G(R_i)\|$ where $R_1, R_2, \cdots, R_k$ are non-overlapping elementary figures in D. Now suppose $|E|<\epsilon$ and $V(G, E)=V(G, R_0^0)$. Define $S(C)=V(G, C)$ where C is either an open set or an elementary figure and let R_n be a sequence of elementary figures such that

$$R_n \subset R_{n+1} \subset E, \qquad S(R_n) \to S(E).$$

Let $R_n' = R_0^0 - R_n$, $E' = R_0^0 - E$, so that

$$S(R_0^0) \geqq S(R_n) + S(R_n') \geqq S(R_n) + \int_{R_n} S'(P)dP.$$

Thus since $R_n' \supset E'$ we have

$$S(R_0^0) \geqq S(R_n) + \int_{E'} S'(P)dP$$

or

$$\int_{E'} S'(P)dP \leqq S(R_0^0) - S(R_n) \to 0.$$

Whence it follows that $S'(P)=0$ almost everywhere on E', and since $|R_0-E'|<\epsilon$ we conclude that $S'(P)=0$ almost everywhere on R_0 which implies $G'(P)=0$ almost everywhere on R_0.

To prove the converse let $\eta>0$ and let $R_1, R_2, \cdots, R_k$ be non-overlapping elementary figures contained in R_0^0 with

$$\left|\sum_{i=1}^{k} R_i\right| > |R_0| - \frac{\epsilon}{2}, \qquad \sum_{i=1}^{k}\|G(R_i)\| \geqq V(G, R_0^0) - \eta.$$

Now define the set E_n as follows: a point P is in E_n if for every cube I containing P with $|I| \leqq 1/n$ it follows that $\|G(I)\|/|I| \leqq \eta$. Thus the set $\lim E_n$ contains all points at which $G'(P)=0$, i.e., almost all points in R_0. Hence $\lim E_n \sum_{i=1}^{k} R_i^0$ contains almost all points of $\sum_{i=1}^{k} R_i^0$. Consequently there is an n_0 and a closed set C contained in $E_{n_0}\sum_{i=1}^{k} R_i^0$ for which $|C| > |R_0| - \epsilon$. Thus $C \subset \sum_{i=1}^{k} R_i^0$ and $\|G(I)\|/|I| \leqq \eta$ for any cube I containing a point of C and having $|I| \leqq 1/n_0$. Let d be the distance from C to the boundary of $\sum_{i=1}^{k} R_i^0$ and $I_1, I_2, \cdots, I_l$ be non-overlapping cubes satisfying the conditions

$$\sum_{i=1}^{l} I_i \supset R_0; \quad |I_i| \leqq 1/n_0, \quad \text{diameter } I_i < d, \qquad (i = 1, 2, \cdots, l).$$

Define δ_j $(j=1, 2, \cdots, l)$ to be 0 or 1 according as I_jC is null or non-null and let $\delta_j' = 1 - \delta_j$. Now $D=D(\epsilon, \eta)$, the complement of C with respect to R_0^0, is open and since R_iI_j $(j=1, 2, \cdots, l)$ is null or a cube if $\delta_j=1$, it follows that

$$\begin{aligned} V(G, R_0^0) - \eta &\leqq \sum_{i=1}^{k} \|G(R_i)\| \leqq \sum_{i=1}^{k}\sum_{j=1}^{l} \|G(R_iI_j)\| \\ &\leqq \sum_{i=1}^{k}\sum_{j=1}^{l} \delta_j \|G(R_iI_j)\| + \sum_{i=1}^{k}\sum_{j=1}^{l} \delta_j' \|G(R_iI_j)\| \\ &\leqq \sum_{i=1}^{k}\sum_{j=1}^{l} \eta |R_iI_j| + V(G, E) \leqq \eta |R_0| + V(G, D). \end{aligned}$$

By defining $E=\sum_{n=1}^{\infty} D(\epsilon/2^n, 1/n)$ the conclusion is immediate.

LEMMA 2. *If $F(R)$ is an additive function of bounded variation defined for elementary figures R contained in a fixed figure R_0 with values in a Banach space satisfying postulates* (A) *and* (B) *then there exist additive functions $\alpha(R)$, $\beta(R)$ such that*

$$F(R) = \alpha(R) + \beta(R),$$

$\alpha(R)$ is an indefinite integral and $\beta(R)$ has a derivative equal to zero almost everywhere on R_0.

Let $F(R)=\sum_{i=1} a_i(R)\phi_i$ and define

$$\alpha(R) = \int_R \left(\sum_{i=1}^{\infty} a_i'(P)\phi_i \right) dP, \qquad \beta(R) = F(R) - \alpha(R).$$

If we write $\beta(R)=\sum_{i=1}^{\infty} b_i(R)\phi_i$ and denote $\sum_{i=1}^{n} b_i(R)\phi_i$ by $\beta_n(R)$ then $\beta_n'(P)=0$ almost everywhere on R_0. Take $\epsilon>0$ and let (Lemma 1) E_n be an open set with $|E_n| < \epsilon/2^n$ such that $V(\beta_n, E_n) = V(\beta_n, R_0^0)$ $(n=1, 2, \cdots)$. Let $E=E_1+E_2+E_3+\cdots$. Now the $|E| < \epsilon$ and $V(\beta_n, E) = V(\beta_n, R_0^0)$. Since $V(\beta_n, R_0^0) \to V(\beta, R_0^0)$ by semi-continuity,* we have

$$V(\beta, E) \geqq V(\beta_n, E) \to V(\beta, R_0^0).$$

Hence $V(\beta, E) = V(\beta, R_0^0)$ so that by Lemma 1, $\beta'(P)=0$ almost everywhere on R_0. This completes the proof of Lemma 2.

Returning to the argument of the theorem itself we see immediately that

* It is well known that the relation $\beta(R) \to \beta(R)$ for $R \subset R_0$ implies $\liminf_n V(\beta_n, R_0) \geqq V(\beta, R_0)$. It is likewise readily seen that $\liminf_n V(\beta_n, R_0^0) \geqq V(\beta, R_0^0)$.

$F(R)$ has a derivative $F'(P)=G(P)$ almost everywhere on R_0. This follows from the fact that $\alpha(R)$, being an indefinite integral, is differentiable with $\alpha'(P)=G(P)$ almost everywhere on R_0. Now if $F(R)$ is absolutely continuous so are the functions $a_i(R)$ and hence

$$\int_R F'(P)dP = \sum_{i=1}^{\infty} \phi_i T_i \int_R F'(P)dP = \sum_{i=1}^{\infty} \phi_i \int_R T_i F'(P)dP\dagger$$

$$= \sum_{i=1}^{\infty} \phi_i \int_R a_i'(P)dP = \sum_{i=1}^{\infty} a_i(R)\cdot\phi_i = F(R)$$

for every elementary figure R in R_0.

3. $L_p(p>1)$ **has the property** (A). Let $\{\phi_i\}$ be the orthonormal sequence of Haar. Schauder‡ has shown that $\{\phi_i\}$ is a base for L_p $(p\geqq 1)$. The sequence $\{\phi_i\}$ also determines the sequence $\{T_i\}$ of linear functionals on L_p by the formula

$$T_i\psi = \int_0^1 \phi_i(t)\psi(t)dt.$$

If $p>1$ this sequence forms a fundamental set in $\bar{L}_p$ (the space conjugate to L_p) in the sense that every point in $\bar{L}_p$ can be approached by finite linear combinations of the elements of the sequence $\{T_i\}$. Now suppose $a_1, a_2, \cdots$ is an arbitrary sequence of real numbers such that $\|x_n\|$ is bounded, where $x_n=\sum_{i=1}^{n} a_i\phi_i$. We have

$$T_i x_n = a_i, \qquad (i \leqq n), \tag{5}$$

and so§ x_n is a weakly convergent sequence in L_p. Since L_p is weakly complete there is a point $x=\sum_{i=1}^{\infty} T_i x\phi_i$ in L_p such that $Tx_n\to Tx$ for every T in $\bar{L}_p$. Now from (5) $a_i=\lim_n T_i x_n = T_i x$ and so $x=\sum_{i=1}^{\infty} a_i\phi_i$ which was to be proved.

4. **Differentiability of functions of bounded variation.** It is the purpose of this paragraph to prove the final assertion in the introduction. Let $f(t)$ be of bounded variation on $(0, 1)$ to X, and let E be the set of functional values of the strictly monotone real function

$$\sigma(t) = t + V(f; 0, t), \qquad (0 \leqq t \leqq 1).$$

The symbol $V(f; a, b)$ stands for the total variation of f on $a\leqq t\leqq b$. Let $\tau(s)$

---

† For the interchange of T_i and $\int_R$ see Garrett Birkhoff, these Transactions, vol. 38 (1935), p. 371.

‡ J. Schauder, *Eine Eigenschaft des Haarschen Orthogonalsystem*, Mathematische Zeitschrift, vol. 28 (1928), pp. 317–320.

§ Banach, loc. cit., p. 133, Theorem 1. This theorem needs to be modified so as to apply to weakly convergent sequences rather than sequences weakly convergent to a point.

on E to $(0, 1)$ be the inverse of $\sigma(t)$ and let $g(s)$ be defined on E by the equation $g(s)=f(\tau(s))$. Now for any two points $s<s'$ in E,

$$\begin{aligned} \|g(s') - g(s)\| &\leqq V(f; \tau(s), \tau(s')) \\ &\leqq \tau(s') - \tau(s) + V(f; \tau(s), \tau(s')) \\ &\leqq \sigma(\tau(s')) - \sigma(\tau(s)) = s' - s. \end{aligned} \tag{6}$$

By first extending the domain of definition of $g(s)$ to $\overline{E}$ (the closure of E) in the natural way and then in a linear fashion on each of the intervals which make up the complement of $\overline{E}$ with respect to the interval $0 \leqq s \leqq 1+V(f; 0, 1)$ it is seen that the extended function satisfies the same Lipschitz condition (6) on the whole of $0 \leqq s \leqq 1+V(f; 0, 1)$. Thus $g(s)$ has a derivative almost everywhere on $(0, 1+V(f; 0, 1))$ and hence almost everywhere on E with respect to E. Now $\tau(s)$ satisfies the Lipschitz condition $|\tau(s')-\tau(s)| \leqq |s'-s|$, and hence if we let E^* be those points of E at which g has a derivative with respect to E we have $m[\tau(E-E^*)]=0$. That is, for almost all t in $(0, 1)$ the point $\sigma(t)$ is in E^*. Thus for almost all t in $(0, 1)$ we have

$$\begin{aligned} \lim_{h\to 0} \frac{f(t+h) - f(t)}{h} &= \lim_{h\to 0} \frac{g(\sigma(t+h)) - g(\sigma(t))}{\sigma(t+h) - \sigma(t)} \lim_{h\to 0} \frac{\sigma(t+h) - \sigma(t)}{h} \\ &= g'(\sigma(t)) \cdot \sigma'(t), \end{aligned}$$

so that $f(t)$ has a derivative at almost all points of $(0, 1)$, and since this derivative is the product of a bounded measurable function and a real summable function it follows that $f'(t)$ is summable.

BROWN UNIVERSITY,
PROVIDENCE, R. I.

THE SPACE OF FUNCTIONS OF BOUNDED VARIATION AND CERTAIN GENERAL SPACES*

BY

C. RAYMOND ADAMS

1. **Introduction.** With usual interpretations of addition, multiplication by a real number, and equality, the space (BV) of functions $x(t)$ of bounded variation on the interval† $(0,1)$ is a vector space.‡ Various writers§ have introduced a metric by defining the norm $\|x\|$ and the distance function (x, y) as follows: $\|x\| = |x(0)| + T_0^1(x)$, where $T_0^1(x)$ designates the total variation of $x(t)$ on $(0, 1)$; $(x, y) = \|x-y\|$. Thus metrized, (BV) is a Banach space (or space of type (B)); so also are its subspaces (CBV), composed of the continuous functions in (BV), and (AC), composed of the absolutely continuous functions. Although with this metric (AC) is a separable space, neither (BV) nor (CBV) is.

Indeed‖ Banach and Mazur, loc. cit., have noted that ("ersichtlich") every set in (BV) whose cardinal number is less than that of the continuum is non-dense in (BV). The reasoning may run as follows: let $x_h(t)$ $(0<h<1)$ be the characteristic function of the point $t=h$; then for $h_1 \neq h_2$ we have $(x_{h_1}, x_{h_2}) = T_0^1(x_{h_1}-x_{h_2}) = 4$, and there exists a continuum of disjoint spheres¶ $K(x_h, 1) \subset K(\theta, 3)$; hence an arbitrary sphere $K(x, r) \subset (BV)$ contains a continuum of disjoint spheres $K(rx_h/3+x, r/3)$. For (CBV), as well as for its subspace (CS) of continuous singular functions, the same property can be demonstrated in a similar way by the aid of some results due to Wiener and

---

* Presented to the Society, April 11, 1936, under a slightly different title; received by the editors March 10, 1936.

† The *interval* $(0, 1)$ is always understood to mean $0 \leqq t \leqq 1$.

‡ In this paper we generally employ the terminology and notation of Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932.

§ These include Hahn, *Über Folgen linearer Operationen*, Monatshefte für Mathematik und Physik, vol. 32 (1922), pp. 3–88; Plessner, *Eine Kennzeichnung der totalstetiger Funktionen*, Crelle's Journal, vol. 160 (1929), pp. 26–32; Hildebrandt, *Linear functional transformations in general spaces*, Bulletin of the American Mathematical Society, vol. 37 (1931), pp. 185–212; Banach and Mazur, *Zur Theorie der linearen Dimension*, Studia Mathematica, vol. 4 (1933), pp. 100–112; and Gergen, *Note on linear operations in functions of bounded variation* (abstract), Bulletin of the American Mathematical Society, vol. 40 (1934), p. 657.

‖ We are indebted to J. A. Clarkson for a suggestion in connection with the contents of this paragraph, which was added August 24, 1936.

¶ In general the zero-element of a space will be designated by θ.

Young.* Perhaps the simplest procedure is to define $x(t)$ on $(0, \frac{1}{2})$ as Wiener and Young's modification (loc. cit., §6) of Cantor's ternary function, on $(\frac{1}{2}, 1)$ as $x(1-t)$, and elsewhere as the (continuous) periodic continuation. Then, allowing $x_h(t)$ to stand for $x(t+h)$, we have $(x_{h_1}, x_{h_2}) = (x_{h_2-h_1}, x) = 2T_0^1(x) = 2$ for $0<h_1<h_2<1$.

One would expect the lack of separability to lead to difficulties in handling certain questions concerning (BV) and (CBV), including that of determining the form of the general linear functional† (i.e., real-valued linear operation) on (BV).

In recent papers by Clarkson, Lewy, and the present writer‡ consideration has been given to a sequence of functions $x_n(t)$ $(n=1, 2, \cdots)$, of bounded variation, subject to the following conditions: (i) $x_n(t)$ tends everywhere on $(0,1)$ to a limit function $x(t)\varepsilon(BV)$, as $n\to\infty$; and (ii) $T_0^1(x_n)\to T_0^1(x)$ [or (ii′) $L_0^1(x_n)\to L_0^1(x)$, where $L_0^1(x)$ denotes the length, in the sense of Peano, of $x(t)$ on $(0, 1)$]. To describe this situation the notation $x_n(t)-v\to x(t)$ $[x_n(t)-l\to x(t)]$ was introduced, the symbol $-v\to[-l\to]$ being read "converges in variation" ["converges in length"]. These studies naturally lead one to contemplate the possibility of defining a metric for (BV) in such a manner that convergence in the metric will have a meaning as close as possible to convergence in variation or to convergence in length. From this standpoint the following four metrics all invite attention:

$$(x, y)_1 = \int_0^1 |x(t) - y(t)|\,dt + |T_0^1(x) - T_0^1(y)|,$$

$$(x, y)_2 = \int_0^1 |x(t) - y(t)|\,dt + |L_0^1(x) - L_0^1(y)|,$$

$$(x, y)_3 = \sup_{0\leq t\leq 1} |x(t) - y(t)| + |T_0^1(x) - T_0^1(y)|,$$

$$(x, y)_4 = \sup_{0\leq t\leq 1} |x(t) - y(t)| + |L_0^1(x) - L_0^1(y)|.$$

Among these metrics the relations $(x, y)_1 \leqq (x, y)_3$, $(x, y)_2 \leqq (x, y)_4$ for every $x, y\varepsilon(BV)$ are apparent. With the metric $(x, y)_3$ $[(x, y)_4]$ each point of the set (DBV) (i.e., each discontinuous function of bounded variation) is the center

* Wiener and Young, *The total variation of* $g(x+h)-g(x)$, these Transactions, vol. 35 (1933), pp. 327–340. An independent proof of the non-separability of (CBV) and (CS) has been constructed by A. P. Morse on the basis of results and methods contained in his paper *On convergence in variation*, to appear shortly in these Transactions.

† Cf. Gergen, loc. cit.; his form for the functional is $f(x)=\int_0^1 K(x, t)dx(t)$.

‡ Adams and Clarkson, *On convergence in variation*, Bulletin of the American Mathematical Society, vol. 40 (1934), pp. 413–417; Adams and Lewy, *On convergence in length*, Duke Mathematical Journal, vol. 1 (1935), pp. 19–26. Hereinafter these papers will be referred to as I and II respectively.

of a sphere containing no points* of (CBV), in marked contrast to the fact that a sequence of continuous functions may converge in variation [length] to a discontinuous function.

We observe that the relation $x_n - v \to x$ implies $(x_n, x)_1 \to 0$; for the functions $x_n(t)$ are uniformly bounded and tend to $x(t)$ everywhere on $(0, 1)$. Conversely, when $x(t)$ is continuous on $(0, 1)$, the relation† $(x_n, x)_1 \to 0$ implies $x_n - v \to x$; this may readily be proved by aid of Theorem 3 of I and the fact that every subsequence of $x_n(t)$ contains a subsequence which converges almost everywhere. If, however, we have $(x_n, x)_1 \to 0$ and $x(t)$ is discontinuous at even a single point, $x_n(t)$ may fail to converge to $x(t)$ at every point in $(0, 1)$; this may be shown by a trivial modification of a familiar type of example illustrating a sequence which converges in the mean without converging anywhere.‡ Thus, although in the space (CBV) the two kinds of convergence are equivalent, in the space (BV) convergence in the metric $(x, y)_1$ is a somewhat weaker condition than convergence in variation.

Precisely the same statements hold for the implications, or lack of the same, between convergence in length and convergence in the metric $(x, y)_2$.

As for the relationship between convergence in the metrics $(x, y)_1$ and $(x, y)_2$, we see that when $x(t)$ is continuous on $(0, 1)$, $(x_n, x)_2 \to 0$ implies $x_n - l \to x$, which by Theorem 1 of II implies $x_n - v \to x$, which implies $(x_n, x)_1 \to 0$. The condition $(x_n, x)_2 \to 0$ fails to imply $(x_n, x)_1 \to 0$, however, when $x(t)$ has even a single discontinuity; this may be seen by slightly altering an example given in §5 of I, merely introducing a suitable jump in the limit function (there called f_0) at $t = 1$. From Theorem 1 of II it is clear that, even in the space (AC), $(x_n, x)_1 \to 0$ does not imply $(x_n, x)_2 \to 0$.

With either of the metrics $(x, y)_3$ and $(x, y)_4$ the spaces (BV), (CBV), and (AC) are not complete and not even locally compact. To verify this one need consider only such a simple example as a sequence of polygonal functions all having the same total variation and the same length on $(0, 1)$ and converging uniformly to $x(t) \equiv 0$.

With the metric $(x, y)_1$, however, every bounded portion of (BV) is compact. For the set of functions $x(t)$ satisfying the condition $(x, \theta)_1 = \int_0^1 |x(t)|\, dt$

* See the corollary to Theorem 4 of I and the corollary to Theorem 2 of II.

† Each continuous function of bounded variation is an element of the metric space (BV), metrized with $(x, y)_1$; on the other hand an element of (DBV) is not a single discontinuous function but a class of such, any two functions in this class being equal almost everywhere and having the same total variation. Here and throughout the remainder of the Introduction, if the sequence $x, x_1, x_2, \cdots$ contains discontinuous functions, it is really non-essential whether one regards the relation $(x_n, x)_1 \to 0$ $[(x_n, x)_2 \to 0]$ as a condition on a sequence of single functions or on a sequence of classes of (metrically equal) functions.

‡ See, for instance, Titchmarsh, *The Theory of Functions*, Oxford, 1932, p. 390, Example (ii).

$+T_0^1(x) \leqq M$ is clearly of uniformly bounded variation and hence is readily seen to be uniformly bounded. By a well known theorem of Helly,* this set contains a sequence $x_n(t)$ converging everywhere to a function $x(t)\varepsilon(BV)$, and by Lebesgue's convergence theorem $x_n(t)$ converges in the mean to $x(t)$. But, the sequence $T_0^1(x_n)$ being bounded, it contains a subsequence $T_0^1(x_{n_i})$ convergent to a limit τ; and by aid of the semi-continuity property,† $\lim\inf T_0^1(x_{n_i}) \geqq T_0^1(x)$, we infer $\tau \geqq T_0^1(x)$. Redefining $x(0)$, if necessary, so that $T_0^1(x)=\tau$, we have $(x_{n_i}, x)_1 \to 0$ and the compactness is proved. The completeness of the space follows at once. Since uniform boundedness of the lengths of a set of functions implies that condition on their total variations,† and since length also enjoys the semi-continuity property,† the same sort of reasoning can be employed to establish the same results for (BV) metrized with $(x, y)_2$.

With either metric $(x, y)_1$ or $(x, y)_2$ the space (BV) is separable; indeed there exists a denumerable subset of (AC) which is dense in (BV). For each metric this assertion may be established as follows. If $x(t)$ is any function in (BV), it can be approximated† in the metric by an inscribed polygonal function $y(t)$. But any such function can obviously be approximated in the metric by a continuous polygonal function $z(t)$ whose "corners" have rational coordinates. Clearly the functions $z(t)$ are a denumerable subset of (AC).

These considerations make it apparent that, of the four metrics in question, the first two are to be preferred above the last two. Of $(x, y)_1$ and $(x, y)_2$, only the former enjoys the homogeneity property $(ax, ay)=|a|\cdot(x, y)$, for every real number a and every x, $y\varepsilon(BV)$, a property whose importance will become amply evident in §2. *Henceforth in this paper the metric* $(x, y)_1$ *will always be employed for* (BV) *unless the contrary is expressly stated.* From an example given in §5 of I it is clear that convergence in this metric, even when we restrict ourselves to the space (AC), is not in general an additive property. We therefore devote §2 to the development of certain elementary properties of spaces which are more weakly conditioned by postulate than are the spaces considered in Banach's book (loc. cit.). Although motivated by our present needs, the discussion contained in §2 may be of some intrinsic interest, dealing as it does with spaces in which the translation property $(x, y)=(x-y, \theta)$ is absent, the relation $\lim(x_n+y_n)=\lim x_n+\lim y_n$ does not in general hold, continuity of an additive and homogeneous operation at one point does not imply continuity everywhere else, and continuity of such an operation over a space does not imply uniform continuity. In §3 we make a few remarks of a

* Helly, *Über lineare Funktionaloperationen*, Sitzungsberichte der Wiener Akademie, vol. 121, IIa (1912), p. 283.

† See §2 of II.

more or less geometric nature concerning the space (BV). In §4 we consider functionals on certain general spaces, and in §5 continuous functionals on the space (BV). So far as we are aware, this is the first instance in which the general form of an additive and uniformly continuous functional is determined for a space not of Banach type.

The interesting question as to the general form of an additive functional which is continuous without being uniformly continuous now arises, apparently for the first time. This problem seems to call for a method of treatment essentially different from that introduced by F. Riesz* in 1909, and at present we have nothing to contribute to its solution.

2. **Certain general spaces.** Consideration of (BV) metrized with the distance function $(x, y)_1$ immediately focusses attention upon the meaning of the term *equality* in vector and metric spaces. The desirability of treating this notion postulationally has been remarked, both in print† and otherwise, by various persons. We do not propose to concern ourselves here with postulate theory and in particular shall not investigate questions of independence, but it is essential that our assumptions be clearly understood. For that reason we begin with definitions of vector and metric spaces, placing special emphasis on the notion of equality in the two cases and making a sharp distinction between vector equality and metric equality.

It may be worthwhile at the outset to state our own viewpoint, which is essentially as follows: for a given space the choice of a distance function imposes a definition of metric equality; with this meaning for equality the space may not be a vector space; nevertheless, with a suitable definition of vector equality, it may become a vector space; there is no logical reason why these two kinds of equality must necessarily be the same.

One‡ might wish to metrize, for example, a space of periodic functions having the same period, in such a way that two functions differing only in phase constant would be indistinguishable (i.e., metrically equal); whereas, when two functions were to be added, a distinction (provided by vector inequality) would have to be made between any two functions not identical. Our viewpoint may sometimes be useful in physics; in fact the "new non-

* F. Riesz, *Sur les opérations fonctionnelles linéaires*, Comptes Rendus, Paris, vol. 149 (1909), pp. 974–977.

† See Michal, *A critique of the postulate systems for abstract vector spaces and abstract Hilbert spaces* (abstract), Bulletin of the American Mathematical Society, vol. 39 (1933), p. 881; Taylor, *A reduced set of postulates for abstract Hilbert space*, ibid., vol. 41 (1935), pp. 439–448, especially footnote on p. 439; and Highberg and Taylor, *An independent set of postulates for abstract linear spaces* (abstract), ibid., vol. 41 (1935), p. 802.

‡ The addition (August 24, 1936) of this paragraph was prompted by Koopman's review, cited below.

relativistic point of view" of Dirac* in quantum mechanics is, to say the least, susceptible of an interpretation bearing a close analogy to ours. He now uses the term "state" to refer to the condition of a dynamical system *at one instant* of time, and considers the aggregate of states of a particular system as constituting a vector space. In this space we understand the element to be a single state, equality between states meaning identity and being analogous to our vector equality; and the law of combination is a "superposition" of two states the result of which is a third state whose relationship to the first two is expressible linearly, with complex numbers for coefficients. At the same time his "further assumption that by superposing a state with itself" one can form "only the original state over again" seems to imply an indistinguishability (between $a\psi$ and ψ) analogous to our metric equality.

DEFINITION 1. A class E of elements is said to constitute a *vector space* when the following postulates are satisfied: (1) to each ordered pair of elements $x, y \in E$ there corresponds an element of E designated by $x+y$ and called the *sum* of x and y; (2) to each real number a and each element $x \in E$ there corresponds an element of E designated by ax and called the *product* of x by a; (3) there exists a relation called *vector equality* and designated by the symbol $=$ such that if x, y is any ordered pair of elements in E, either x bears this relation to y (in which case we write $x=y$) or it does not (in which case we write $x \neq y$); (4) the relation $=$ is *reflexive* (i.e., $x=x$ for every $x \in E$), *symmetric* (i.e., $x=y$ implies $y=x$), and *transitive* (i.e., $x=y$ and $y=z$ together imply $x=z$); and (5) with respect to this relation $=$, the sum and product fulfill the conditions $x+y=y+x$, $x+(y+z)=(x+y)+z$, $x+y=x+z$ implies $y=z$, $a(x+y)=ax+ay$, $(a+b)x=ax+bx$, $a(bx)=(ab)x$, $1x=x$, where a, b are any real numbers and x, y, z any elements in E.

We recall that in a vector space E there exists a zero-element θ (unique in the sense determined by vector equality) having the properties: $x+\theta=x$ for every $x \in E$; $a\theta=\theta$ for every real number a.

DEFINITION 2. A class E of elements is said to constitute a *metric space* when the following postulates are satisfied: (1) to each ordered pair of elements $x, y \in E$ there corresponds a real number called the *distance* from x to y and designated by (x, y); (2) there exists a relation called *metric equality* (or *metric equivalence*) and designated by the symbol $\cong$ such that if x, y is any ordered pair of elements in E, either x stands in this relation to y (we then

* Dirac, *The Principles of Quantum Mechanics*, 2d edition, Oxford, 1935, especially Chapters I, II. In the language (punctuation included) of his reviewer (Koopman, Bulletin of the American Mathematical Society, vol. 42 (1936), p. 472), "The point of view is to regard the aggregate of states of a system somehow as forming a linear vector space, the elements of which both are and are not the states (there is an indeterminate phase factor!), and ···."

write $x \cong y$) or it does not (we then write x *not* $\cong y$); (3) the relation $\cong$ is *reflexive*, *symmetric*, and *transitive*; and (4) with respect to this relation $\cong$, the distance function satisfies the conditions $(x, y) = 0$ for $x \cong y$, $(x, y) > 0$ for x *not* $\cong y$, $(x, y) = (y, x)$, $(x, z) \leqq (x, y) + (y, z)$ (the *triangle postulate*), for any elements $x, y, z \in E$.

We observe that in a metric space E, for any $x, y, z \in E$,

$$(2.1) \qquad x \cong y \quad \textit{implies} \quad (x, z) = (y, z).$$

This follows at once since we have

$$(x, z) \leqq (x, y) + (y, z) = (y, z) \leqq (y, x) + (x, z) = (x, z).$$

The statement that $U(x)$ is an operation defined on a space E with counter-domain in a space E' (or more briefly, *on E to E'*) is commonly understood, and we shall so understand it, to mean that to each element $x \in E$ there corresponds an element $U(x) \in E'$. The following postulate, which appears to have been tacitly assumed by some writers, we impose explicitly.

POSTULATE 1. An operation $U(x)$ shall be said to be *defined* on a vector space E to a like space E' only if $x = y \in E$ implies $U(x) = U(y) \in E'$.

For our present purposes it seems desirable (although it may not always be so) to impose also

POSTULATE 2. In a vector metric space E the relation $x = y$ shall imply $x \cong y$, for $x, y \in E$.

DEFINITION 3. In a vector metric space we define the *norm* of any element x, written $\|x\|$, as (x, θ).

DEFINITION 4. An operation $U(x)$ on a metric space E to a like space shall be called *continuous at* x when for $x_n, x \in E$

$$(x_n, x) \to 0 \quad \text{implies} \quad (U(x_n), U(x)) \to 0;$$

continuous on E when it is continuous at every point of E; and *uniformly continuous on* E when for every $x_n, y_n \in E$

$$(x_n, y_n) \to 0 \quad \text{implies} \quad (U(x_n), U(y_n)) \to 0.$$

THEOREM 2.1.* *Let $U(x)$ be an operation on a metric space E to a like space. If $U(x)$ is continuous at a point $x \in E$ and $x \cong y$, a necessary and sufficient condition that $U(x)$ be continuous at y is $U(x) \cong U(y)$.*

* If operations not everywhere continuous were to be considered, one would probably find it desirable to replace this theorem by the following analogue of Postulate 1: an operation $U(x)$ shall be said to be *defined* on a metric space E to a like space E' only if $x \cong y \in E$ implies $U(x) \cong U(y) \in E'$.

Proof. For the sufficiency, let x_n be any sequence such that $(x_n, y)\to 0$; then $(x_n, x)\to 0$, which implies $(U(x_n), U(x))=(U(x_n), U(y))\to 0$. For the necessity, we have both $(x_n, y)\to 0$ and $(x_n, x)\to 0$; hence both $(U(x_n), U(y))\to 0$ and $(U(x_n), U(x))\to 0$, and $(U(x), U(y))=0$ follows from the triangle postulate.

The *space of real numbers*, with the usual interpretations, is a Banach space; we designate it by (A).

DEFINITION 5. A vector metric space E in which the metric has the additional property†

$$(2.2) \qquad (ax, ay) = |a| \cdot (x, y) \qquad \text{for every } a\varepsilon(A);\ \text{every } x, y\varepsilon E,$$

shall be called a *space of type* (α^*).

It may be observed that in a space E of type (α^*), the relation $ax\cong y$ $(0\neq a\varepsilon(A); x, y\varepsilon E)$ implies $x\cong (1/a)y$. Addition of metric equalities is not in general permissible, however, even when the space is more strongly conditioned (as it will be later when spaces of types (α^{**}) and (α) are introduced).

DEFINITION 6. An operation $U(x)$ on a vector metric space E to a like space shall be called *strongly additive* when

$$U(x + y) = U(x) + U(y) \qquad \text{for every } x, y\varepsilon E;$$

weakly additive when

$$U(x + y) \cong U(x) + U(y) \qquad \text{for every } x, y\varepsilon E;$$

strongly homogeneous when

$$U(ax) = aU(x) \qquad \text{for every } a\varepsilon(A), \text{ every } x\varepsilon E;$$

and *weakly homogeneous* when

$$U(ax) \cong aU(x) \qquad \text{for every } a\varepsilon(A), \text{ every } x\varepsilon E.$$

THEOREM 2.2. *Let $U(x)$ be an operation on a space E of type (α^*) to a like space.‡ If $U(x)$ is weakly homogeneous and uniformly continuous on E, it satisfies a Lipschitz condition*

$$(2.3) \qquad (U(x), U(y)) \leqq M\cdot(x, y) \qquad \textit{for every } x, y\varepsilon E.$$

Conversely, it is clear that if $U(x)$ satisfies (2.3), it is uniformly continuous on E.

Proof. In the absence of such a constant M, there would exist sequences

† The independence of this property is shown by (BV) metrized with $(x, y)_2$.

‡ It may be noted that the following less restrictive hypotheses on the spaces are sufficient for the validity of this theorem: each space to be one in which multiplication of an element by a real number is possible (i.e., merely property (2), Definition 1, of a vector space), and to be a metric space in which the metric has the homogeneity property (2.2).

$x_n, y_n \varepsilon E$ such that $(U(x_n), U(y_n)) > n \cdot (x_n, y_n)$, with $(x_n, y_n) > 0$ by Theorem 2.1. Designating by X_n, Y_n respectively the elements

$$\frac{1}{n(x_n, y_n)} x_n, \qquad \frac{1}{n(x_n, y_n)} y_n \varepsilon E,$$

we have by (2.2)

$$(X_n, Y_n) = \frac{1}{n} \to 0.$$

Since $U(x)$ is uniformly continuous, this relation must imply $(U(X_n), U(Y_n)) \to 0$. But, $U(x)$ being weakly homogeneous, we have by (2.1) for all n,

$$(U(X_n), U(Y_n)) = \left(\frac{1}{n(x_n, y_n)} U(x_n), \frac{1}{n(x_n, y_n)} U(y_n)\right)$$
$$= \frac{1}{n(x_n, y_n)} (U(x_n), U(y_n)) > 1.$$

From this contradiction follows condition (2.3).

THEOREM 2.3. *Let $U(x)$ be an operation on a space E of type (α^*) to a like space. If $U(x)$ is strongly additive and weakly homogeneous on E and continuous at θ, it satisfies a Lipschitz condition at the point θ,*

$$\|U(x)\| \leqq M_\theta \cdot \|x\| \qquad \text{for every } x \varepsilon E. \tag{2.4}$$

Conversely, it is clear that if $U(x)$ satisfies (2.4), it is continuous at θ.

Proof. From the strong additivity of $U(x)$ we have $U(x) = U(x+\theta) = U(x) + U(\theta)$; hence $U(\theta) = \theta'$, the zero-element of the counter-domain. We may then employ the same kind of argument as in the proof of Theorem 2.2 to obtain the condition (2.4).

COROLLARY. *Let $U(x)$ be an operation on a space E of type (α^*) to a space of type (F) (sense of Banach). If $U(x)$ is additive and homogeneous on E and continuous at θ, it satisfies the condition*

$$(U(x), U(y)) \leqq M_\theta \cdot \|x - y\| \qquad \text{for every } x, y \varepsilon E.$$

DEFINITION 7. Let $U(x)$ be an operation on a space E of type (α^*) to a like space. If $U(x)$ satisfies a Lipschitz condition (2.3), we designate by $\text{mod}_E\, U$, the *modulus of the operation* $U(x)$ on E, the smallest number M for which (2.3) holds. If $U(x)$ satisfies the condition (2.4), we designate by $\|U\|_E$, the *norm of the operation* $U(x)$ on E, the smallest number M_θ for which (2.4) is fulfilled.

If $U(x)$ is weakly homogeneous, it is clear from (2.1) and (2.2) that for every $a \in (A)$ and every $x, y \in E$ we have

$$\frac{(U(ax), U(ay))}{(ax, ay)} = \frac{(aU(x), aU(y))}{|a| \cdot (x, y)} = \frac{(U(x), U(y))}{(x, y)},$$

whence

$$\underset{E}{\operatorname{mod}} U = \sup_{x, y \in E, (x, y)=1} (U(x), U(y)), \tag{2.5}$$

$$\|U\|_E = \sup_{x \in E, \|x\|=1} \|U(x)\|. \tag{2.6}$$

We clearly have $\|U\|_E \leqq \operatorname{mod}_E U$, whether both norm and modulus exist finite or otherwise.

DEFINITION 8. A space E of type (α^*) in which the metric has the additional property

$$(a_n x, x) \to 0 \tag{2.7}$$

for every real sequence $a_n \to 1$ and every $x \in E$, shall be called a *space of type* (α^{**}).

The following properties of this type of space may be noted: for $a_n, a \in (A)$ we have

$$a_n \to 0 \text{ implies } \|a_n x\| = |a_n| \cdot \|x\| \to 0 \qquad \text{for every } x \in E; \tag{2.8}$$

$$(x_n, x) \to 0 \text{ implies } (ax_n, ax) = |a| \cdot (x_n, x) \to 0 \qquad \text{for every } a; \tag{2.9}$$

$$a_n \to a \neq 0 \text{ implies } (a_n x, ax) = |a| \cdot \left(\frac{a_n}{a} x, x\right) \to 0 \qquad \text{for every } x \in E; \tag{2.10}$$

whence, by aid of the triangle postulate, it may readily be proved that for $a_m, a, b_n, b \in (A)$ and $x_p, x, y_q, y \in E$, the relations $a_m \to a$, $b_n \to b$, $(x_p, x) \to 0$, $(y_q, y) \to 0$ imply

$$(a_m x_p, b_n y_q) \to (ax, by) \qquad \text{as } m, n, p, q \to \infty.$$

THEOREM 2.4. *Let $U(x)$ be an operation on a space E of type (α^{**}) to a like space. If $U(x)$ is strongly additive [weakly additive and $U(-x) \cong -U(x)$ for every $x \in E$] and continuous on E, it is weakly homogeneous.*

Proof. From the hypotheses it follows easily that $U(rx) = rU(x)$ $[U(rx) \cong rU(x)]$ for every rational r [every rational $r \neq 0$]. Let a be any irrational number and r_n a sequence of rationals converging to a. By (2.10) we have, for every $x \in E$, $(r_n x, ax) \to 0$; in view of the continuity this implies $(U(r_n x), U(ax)) \to 0$, or $(r_n U(x), U(ax)) \to 0$. But since the counter-domain has

property (2.7), $r_n \to a$ also implies $(r_n U(x), aU(x)) \to 0$. From the triangle postulate we conclude $U(ax) \cong aU(x)$. In a similar manner continuity enables us to infer from the bracketed hypotheses $U(\theta) \cong \theta'$, which completes the proof.

DEFINITION 9. A space E of type (α^*) in which the metric has the additional property $(x, y) \leqq \|x-y\|$ for every $x, y \in E$ shall be called a *space of type* (α).

We observe that every vector metric space has the property $|\,\|x\| - \|y\|\,| \leqq (x, y)$; hence a space of type (α) satisfies the condition $|\,\|x\| - \|y\|\,| \leqq \|x-y\|$, which implies

$$\|x + y\| \leqq \|x\| + \|y\|. \tag{2.11}$$

A space of type (α) is also of type (α^{**}).

Remarks. (I) Every "normed vector space" (sense of Banach) is of type (α). (II) The space (BV) becomes a vector space when we interpret addition and multiplication by a constant in the usual way and $x = y$ as meaning that the functions $x(t)$ and $y(t)$ are identical on the interval $0 \leqq t \leqq 1$. Upon introducing the distance function $(x, y)_1$ and interpreting $x \cong y$ as meaning that the functions $x(t)$ and $y(t)$ are equal almost everywhere and have the same total variation on $0 \leqq t \leqq 1$, it becomes a vector metric space, in fact a space of type (α). Henceforth these interpretations are always to be understood when we speak of "the space (BV)." It is clear that (BV) is not a "normed vector space." (III) Every Banach space (or space of type (B)) is a complete space of type (α). (IV) The class of Banach spaces is the meet of the class of Fréchet spaces (or spaces of type (F)) and the class of complete spaces of type (α^*) in which $=$ and $\cong$ have the same meaning. (V) A space [complete space] of type (α^*) is a "normed vector space" [Banach space] if and only if the metric has the additional property $(x, y) = \|x-y\|$ and there is no distinction between metric and vector equality.

3. **Some properties of the space** (BV). Let x, y be any two points in (BV) with $x \neq y$; then the set of points

$$z = hx + (1 - h)y \qquad (0 \leqq h \leqq 1)$$

constitutes, by definition, the *segment* xy. If $h_0, h_1, \cdots, h_n$ is any finite set of values of h satisfying the condition $0 = h_0 < h_1 < \cdots < h_n = 1$ and the corresponding z's are denoted by corresponding subscripts, we define the *length of the segment* xy as the least upper bound of $\sum_{i=1}^{n} (z_{i-1}, z_i)$ for all such sets. With this terminology one may easily verify the following statements. A segment is a continuous curve; whence it follows that (BV) is a connected space.

Each segment xy has a finite positive length which does not exceed $\int_0^1 |x-y|\,dt+T_0^1(x-y)$. This bound for the length of a segment is attainable. A segment may have a continuum of multiple points (multiple in the sense determined by metric equality). When x, y are on the surface of a sphere, the segment xy may contain points outside the sphere.

4. **Linear functionals on a space of type (α).** By the term functional we mean an operation $U(x)$ on a space E to the space (A) of real numbers. Noting that for a functional the qualification of additivity and homogeneity by "strong" and "weak" loses its significance, we establish two theorems as follows.

THEOREM 4.1. *Given a space G of type (α) embedded in a space E of the same type, and a functional $f(x)$ additive and homogeneous on G and continuous on G at θ; then there exists a functional $F(x)$, additive and homogeneous on E and continuous on E at θ, which satisfies the conditions $F(x)=f(x)$ for every $x\varepsilon G$ and $\|F\|_E=\|f\|_G$.*

Proof. For $x\varepsilon E$ we set $p(x)=\|f\|_G\cdot\|x\|$ in the Hahn-Banach theorem on extension of functionals (Theorem 1, Banach, loc. cit., pp. 27–28). Then by (2.11) we have for x, $y\varepsilon E$, $p(x+y)\leqq p(x)+p(y)$; and by (2.2) for $x\varepsilon E$ and $t\geqq 0$, $p(tx)=tp(x)$. E and G being vector spaces, $f(x)$ additive and homogeneous on G, and $f(x)\leqq p(x)$ on G, we conclude from the Hahn-Banach theorem the existence on E of an additive and homogeneous functional $F(x)$ coinciding with $f(x)$ on G and having the property $F(x)\leqq p(x)$ for every $x\varepsilon E$. It remains only to show that $F(x)$ is continuous on E at θ and $\|F\|_E=\|f\|_G$. For any $x\varepsilon E$ for which $F(x)$ is $\geqq 0$ we have $F(x)\leqq p(x)=\|f\|_G\cdot\|x\|$; for any $x\varepsilon E$ for which $F(x)$ is <0 the homogeneity of $F(x)$ gives us $-F(x)=F(-x)\leqq p(-x)=\|f\|_G\cdot\|x\|$. Hence $\|F\|_E\leqq\|f\|_G$, and the inequality is ruled out by the fact that $F(x)=f(x)$ on G.

THEOREM 4.2. *For every $x_0\neq\theta$ in a space E of type (α) there exists an additive and homogeneous functional $F(x)$, defined on E and continuous at θ, which is uniformly continuous on a subspace G of type (α) containing x_0 and which satisfies the conditions $F(x_0)=\|x_0\|$ and $\|F\|_E=1$.*

Proof. It suffices to designate by G in Theorem 4.1 the set of elements ax_0, $a\varepsilon(A)$, and to set $f(ax_0)=a\cdot\|x_0\|$.

It is natural to raise the question as to whether Theorem 4.1 is the best that can be obtained; in particular, whether the following might be true: given an additive and uniformly continuous functional $f(x)$ defined on a space G of type (α) embedded in a space E of the same type, there exists an additive and uniformly continuous functional $F(x)$ defined on E, with

$F(x)=f(x)$ for $x\epsilon G$ (and with or without an increase in the modulus). If this were so, it would imply that for every $x\neq\theta$ in a space E of type (α) there exists an additive and uniformly continuous functional $F(x)$ defined on E, with $F(x_0)=\|x_0\|$. But (BV) is a space of type (α), and by Theorem 5.2 every additive and uniformly continuous functional on (BV) vanishes at x_0, where $x_0(t)=1$ for $t=0$ and is zero otherwise, although $\|x_0\|=1$. Hence the suggested extension is not always possible.

5. **Functionals on the space (BV).** We first establish

THEOREM 5.1. *If $\alpha(t)$ is any summable function, the integral*

$$\int_0^1 x(t)\alpha(t)\,dt \tag{5.1}$$

defines on the space (BV) a functional $f(x)$ which is additive, continuous, and homogeneous. The functionals generated by $\alpha_1(t)$ and $\alpha_2(t)$ are identical if and only if $\alpha_1(t)$ and $\alpha_2(t)$ are essentially the same on $(0, 1)$. The functional $f(x)$ is uniformly continuous on (BV) if and only if $\alpha(t)$ is essentially bounded on $(0, 1)$; in that case we have* $\operatorname{mod}_{(BV)} f=\operatorname{ess\,sup}_{0\leqq t\leqq 1}|\alpha(t)|$.

Proof. The existence, additivity, and homogeneity of $f(x)$ are apparent. To prove the continuity, let x be any point $\epsilon(BV)$ and x_n any sequence such that $(x_n, x)\to 0$. There exists a number K' for which we have for all n

$$(x_n, x)=\int_0^1 |x_n-x|\,dt+|T_0^1(x_n)-T_0^1(x)|\leqq K';$$

hence for each n there is at least one point $\bar{t}_n$ such that

$$|x_n(\bar{t}_n)-x(\bar{t}_n)|\leqq K'.$$

But for each n and all t we have

$$|x_n(t)-x_n(\bar{t}_n)|\leqq T_0^1(x_n),\quad |x(t)-x(\bar{t}_n)|\leqq T_0^1(x),\quad T_0^1(x_n)\leqq T_0^1(x)+K',$$

whence for all n and t

$$|x_n(t)-x(t)|\leqq 2T_0^1(x)+2K'=K.$$

Moreover, given $\epsilon>0$ there exists $\delta>0$ such that $\int_E K|\alpha|\,dt<\epsilon$ whenever $mE<\delta$. Choosing N so that $mE_t[|\alpha(t)|>N]<\delta$, we have

$$|f(x_n)-f(x)|=|f(x_n-x)|\leqq\left[\int_E+\int_{CE}\right]|x_n-x|\cdot|\alpha|\,dt$$

$$\leqq\epsilon+N\int_0^1|x_n-x|\,dt,$$

which is $<2\epsilon$ for n sufficiently large.

* That is, $\alpha_1(t)=\alpha_2(t)$ almost everywhere.

It is clear that if $\alpha_1(t)$ and $\alpha_2(t)$ are essentially the same, they generate the same functional. On the other hand if the same functional is generated by $\alpha_1(t)$ and $\alpha_2(t)$, we may choose $x(t)=1$ for $0\leqq t\leqq y$, $=0$ for $y<t\leqq 1$ $(0\leqq y\leqq 1)$ and obtain $\int_0^y \alpha_1(t)dt=\int_0^y \alpha_2(t)dt$ for all y. The essential equality of $\alpha_1(t)$ and $\alpha_2(t)$ then follows at once. This fact, in conjunction with the proof* of the following Theorem 5.2, shows that a uniformly continuous functional $f(x)$ cannot be generated by an $\alpha(t)$ which is not essentially bounded.

That $f(x)$ is uniformly continuous when $\alpha(t)$ is essentially bounded follows at once from (2.5); we have

$$\operatorname*{mod}_{(BV)} f = \sup_{x,\, y\in(BV),\ (x,\,y)=1} |f(x)-f(y)| \leqq \operatorname*{ess\,sup}_{0\leqq t\leqq 1} |\alpha(t)| = B.$$

To show that in this case $\operatorname{mod}_{(BV)} f=B$, we may proceed as follows. Let $\epsilon(>0)$ be arbitrarily small and set

$$mE_1[\alpha(t) > B-\epsilon] = k_1, \qquad mE_2[\alpha(t) < \epsilon - B] = k_2.$$

Then either k_1 or k_2 is >0; the cases being symmetrical, we consider that in which $k_1>0$. Choose δ to satisfy the condition $0<\delta<k_1\epsilon$ and inclose E_1 in an open set O_1 such that $k_1\leqq mO_1\leqq k_1+\delta$. From O_1 select a finite number of open intervals constituting a set O_2 such that $mO_2\geqq mO_1-\delta\geqq k_1-\delta$, and define $\phi(t)=1/k_1$ for $t\in O_2$, $=0$ elsewhere. Then $\phi(t)$ is in (BV) and we have

$$\int_0^1 \phi(t)\alpha(t)dt \geqq (k_1-\delta)\frac{1}{k_1}(B-\epsilon) - \delta\frac{1}{k_1}B \geqq B-\epsilon-2B\epsilon,$$

$$\int_0^1 \phi(t)dt \leqq \frac{1}{k_1}(k_1+\delta) \leqq 1+\epsilon.$$

Finally we set $x(t)=-y(t)=\phi(t)/2$ for all t in $(0,1)$; then $T_0^1(x)=T_0^1(y)$ and we have

$$|f(x)-f(y)| = \left|\int_0^1 (x-y)\alpha dt\right| = \int_0^1 \phi(t)\alpha(t)dt \geqq B-\epsilon-2B\epsilon,$$

$$(x,\,y) = \int_0^1 |x-y|\,dt = \int_0^1 \phi(t)dt \leqq 1+\epsilon,$$

whence

$$|f(x)-f(y)|/(x,\,y) \geqq (B-\epsilon-2B\epsilon)/(1+\epsilon).$$

Since the latter quantity is arbitrarily close to B, we conclude that $\operatorname{mod}_{(BV)} f=B$.

* See the italicized statement there.

THEOREM 5.2. *Every additive and uniformly continuous functional defined on the space* (BV) *can be expressed in the form* (5.1), *with* $\alpha(t)$ *summable, bounded, and independent of* $x(t)$; *moreover we have* $\|f\|_{(BV)} \leqq \int_0^1 |\alpha(t)|\,dt \leqq \mathrm{mod}_{(BV)}\, f = B$, *where* $B = \operatorname{ess\,sup}_{0 \leqq t \leqq 1} |\alpha(t)|$.

Proof. Let $f(x)$ be any additive and uniformly continuous functional defined on (BV). We set

$$\xi_t(u) = \begin{cases} 1 & \text{for } 0 \leqq u \leqq t, \\ 0 & \text{for } t < u \leqq 1, \end{cases}$$

and

$$f(\xi_t) = g(t).$$

For $0 \leqq t_1 < t_2 \leqq 1$ we have

$$\begin{aligned} |\, g(t_1) - g(t_2)\,| &= (f(\xi_{t_1}), f(\xi_{t_2})) \\ &\leqq M \cdot (\xi_{t_1}, \xi_{t_2}) = M \int_0^1 (\xi_{t_2} - \xi_{t_1})\,dt + M\,|\, T_0^1(\xi_{t_2}) - T_0^1(\xi_{t_1})\,|, \end{aligned} \tag{5.2}$$

where $M = \mathrm{mod}_{(BV)}\, f$. For $t_2 < 1$ the second term on the right vanishes, and we have $|\, g(t_1) - g(t_2)\,| \leqq M(t_2 - t_1)$; i.e., $g(t)$ satisfies a Lipschitz condition on the open interval $0 \leqq t < 1$ and is absolutely continuous there. We now define

$$\bar{g}(t) = \begin{cases} g(t) & \text{for } 0 \leqq t < 1, \\ g(1-0) & \text{for } t = 1; \end{cases}$$

the existence of the integral $\int_0^1 x(t)\,d\bar{g}$ is then assured for every $x \in (BV)$.

Let $x(t)$ be any function in (BV) vanishing at $t = 0$. In view of a remark in §2 of II, there exists a sequence of polygonal functions $p_n(t)$ $(n = 1, 2, 3, \cdots)$ inscribed in $x(t)$ and converging in variation to $x(t)$; for each n the "corners" of $p_n(t)$ are determined by a set S_n of points

$$0 = t_{0,n} < t_{1,n} < t_{2,n} < \cdots < t_{r(n),n} = 1$$

with $S_n \subset S_{n+1}$ for every n and $S = \sum_{n=1}^{\infty} S_n$ everywhere dense in $(0, 1)$. We define a step-function $z_n(t)$ associated with $p_n(t)$ as follows:

$$\begin{aligned} z_n(t) &= x(t_{i,n}) \text{ for } t_{i-1,n} < t \leqq t_{i,n}, \text{ all } i \text{ and } n; \\ z_n(0) &= 0. \end{aligned}$$

Then it is clear that $T_0^1(z_n) = T_0^1(p_n)$, whence $T_0^1(z_n) \to T_0^1(x)$ as $n \to \infty$. Moreover, since $x(t)$ is Riemann integrable on $(0, 1)$, $z_n(t)$ converges in the mean to $x(t)$. Hence we have $(z_n, x) \to 0$, which implies $f(z_n) \to f(x)$.

But $z_n(u)$ may be expressed in the form

$$z_n(u) = \sum_i x(t_{i,n})[\xi_{t_{i,n}}(u) - \xi_{t_{i-1,n}}(u)],$$

therefore, since $f(x)$ is additive, we must have

$$f(z_n) = \sum_i x(t_{i,n})[g(t_{i,n}) - g(t_{i-1,n})]$$
$$= \sum_i x(t_{i,n})[\bar{g}(t_{i,n}) - \bar{g}(t_{i-1,n})] + x(1)[g(1) - \bar{g}(1)].$$

Under the conditions in force we infer

$$f(x) = \lim_{n\to\infty} f(z_n) = \int_0^1 x(t)d\bar{g} + x(1)[g(1) - \bar{g}(1)], \tag{5.3}$$

which is the form of the functional for every $x(t)$ vanishing at $t=0$.

Any function $x(t)$ not vanishing at $t=0$ can be expressed as the sum $x_1(t)+x_2(t)$ where

$$x_1(t) = \begin{cases} x(t) & \text{for } 0 < t \leqq 1, \\ 0 & \text{for } t = 0, \end{cases} \qquad x_2(t) = \begin{cases} x(0) & \text{for } t = 0, \\ 0 & \text{for } 0 < t \leqq 1, \end{cases}$$

and, $f(x)$ being additive, we must have $f(x)=f(x_1)+f(x_2)$. But $f(x)$ being continuous, it must assume the same value at any two points x_2, x_3 such that $x_2 \cong x_3$. Choosing x_3 as follows:

$$x_3(t) = \begin{cases} [x(0)]/2 & \text{for } t = 1/2, \\ 0 & \text{otherwise}, \end{cases}$$

we clearly have $x_2 \cong x_3$, while $f(x_3)$ is given by (5.3) and has the value zero. Hence $f(x)$ is given by (5.3) even when $x(0) \neq 0$.

Next consider the functions

$$x_4(t) = \begin{cases} 0 & \text{for } 0 \leqq t < 1, \\ 1 & \text{for } t = 1, \end{cases} \qquad x_5(t) = \begin{cases} 0 & \text{for } 0 \leqq t < 1, \\ -1 & \text{for } t = 1. \end{cases}$$

Since $x_4 \cong x_5$, we must have $f(x_4)=f(x_5)$; hence by (5.3) we have $g(1)-\bar{g}(1) = -[g(1)-\bar{g}(1)]$, whence $g(1)=\bar{g}(1)$ and (5.3) reduces to

$$f(x) = \int_0^1 x(t)dg \qquad \text{for all } x \in (BV), \tag{5.4}$$

where $g(t)$ is absolutely continuous on $0 \leqq t \leqq 1$. *The form* (5.1) *for* $f(x)$ *is an immediate consequence, with* $\alpha(t)$ *bounded*, since in fact $g(t)$ satisfies a Lipschitz condition on $0 \leqq t \leqq 1$. The evaluation of $\mathrm{mod}_{(BV)} f$ has been made in Theorem 5.1.

We now proceed to obtain some estimates for $\|f\|_{(BV)}$. In the first place, from (5.4) we have $|f(x)| \leq [\sup_{0\leq t\leq 1}|x(t)|]T_0^1(g)$ and

$$\frac{|f(x)|}{\|x\|} \leqq \frac{\sup|x|}{\int_0^1|x|+T_0^1(x)}\,T_0^1(g) \leqq T_0^1(g),$$

from which follows the bound asserted in the theorem. Secondly, for $0\leqq t_1<t_2\leqq 1$, we have

$$\begin{aligned}|g(t_1)-g(t_2)| &= |f(\xi_{t_1}-\xi_{t_2})| \leqq \|f\|_{(BV)}\cdot\|\xi_{t_1}-\xi_{t_2}\| \\ &= \|f\|_{(BV)}\cdot\left[\int_0^1(\xi_{t_2}-\xi_{t_1})dt + T_0^1(\xi_{t_2}-\xi_{t_1})\right],\end{aligned}$$

which implies

$$\tfrac{1}{3}\operatorname*{osc}_{0\leq t\leq 1} g(t) \leqq \|f\|_{(BV)}.$$

Thirdly, we observe that for $x(t)\equiv 1$ on $(0,1)$ we have

$$f(x)=f(\xi_1)=g(1)=\int_0^1 dg = g(1)-g(0),$$

which implies $g(0)=0$. Finally, one may determine estimates for $\|f\|_{(BV)}$, which may or may not be closer, by making use of the relation

$$\int_0^1 x(t)dg = x(1)g(1)-\int_0^1 g(t)dx.$$

From this we have

$$\frac{|f(x)|}{\|x\|}=\frac{|f(x)|}{\int_0^1|x|+T_0^1(x)} \leqq |g(1)|\frac{|x(1)|}{\int_0^1|x|+T_0^1(x)}+\max_{0\leq t\leq 1}|g(t)|,$$

whence follows

$$\|f\|_{(BV)} \leqq |g(1)|+\max_{0\leq t\leq 1}|g(t)|. \tag{5.5}$$

At the same time from the relations

$$\begin{aligned}|g(1)| &= \|f(\xi_1)\| \leqq \|f\|_{(BV)}\cdot\|\xi_1\| = \|f\|_{(BV)}, \\ |g(t)| &= \|f(\xi_t)\| \leqq \|f\|_{(BV)}\cdot\|\xi_t\| < 2\|f\|_{(BV)},\end{aligned}$$

we obtain

$$\max\left[|g(1)|,\ \tfrac{1}{2}\max_{0\leq t\leq 1}|g(t)|\right] \leqq \|f\|_{(BV)};$$

and it may be noted that this implies $\|f\|_{(BV)} \geqq K/3$, where K stands for the right-hand member of (5.5).

THEOREM 5.3. *Every additive and uniformly continuous functional $f(x)$ defined on the space (CBV) $[(AC)]$ can be expressed in the form* (5.1) *where $\alpha(t)$ is summable, bounded, and independent of $x(t)$, and* mod $f=B$ *as before.*

Proof. It has already been observed in §1 that (CBV) is dense in (BV). Let x be any point in $(BV)-(CBV)$ and x_n any sequence $\epsilon(CBV)$ such that $(x_n, x)\to 0$. If the functional $f(x)$ is uniformly continuous on (CBV), $\lim f(x_n)$ exists and we may define $f(x)$ as this limit. Moreover this limit is independent of the sequence x_n employed and $f(x)$ thus extended over (BV) is uniformly continuous over (BV). That the extended functional is additive on (BV) may readily be seen as follows. Let x and y be any two points of (BV) at least one of which is not in (CBV). A sequence of sets S_n in (0,1) can be so determined that the polygonal functions $x_n(t)$, $y_n(t)$, and $z_n(t)$ respectively inscribed in x, y, and $x+y$, whose "corners" are determined by S_n, converge in our metric to x, y, and $x+y$. We then have

$$z_n(t) = x_n(t) + y_n(t) \text{ for all } n, t; \qquad f(z_n) = f(x_n) + f(y_n),$$
$$\lim f(x_n) = f(x), \quad \lim f(y_n) = f(y), \quad \lim f(z_n) = f(x+y),$$

whence $f(x+y)=f(x)+f(y)$. That mod $f=B$ can be shown essentially as in the proof of Theorem 5.1, by appropriately smoothing off the function $\phi(t)$ (as it was designated there) to make it continuous. A similar proof will suffice for the space (AC).

BROWN UNIVERSITY,
PROVIDENCE, R. I.

ON IDEALS IN A QUATERNION ALGEBRA AND THE REPRESENTATION OF INTEGERS BY HERMITIAN FORMS*

BY

CLAIBORNE G. LATIMER

1. **Introduction.** In another paper,† it was shown that there is a one-to-one correspondence between certain classes of ideals, called regular classes, in a quaternion ring $\mathfrak{G}$, and certain classes of binary Hermitian forms. A binary Hermitian form is in one of these classes if and only if (a) its coefficients are in the set G of all integral numbers in a certain quadratic field of discriminant Δ, (b) the variables range over G, (c) its determinant is a certain integer α, (d) it represents positive integers.

Let f be such a form and let m be a positive integer. We shall show that f represents m if and only if there is an ideal of norm m in a certain class $\mathfrak{C}$ of ideals in $\mathfrak{G}$. $\mathfrak{C}$ is uniquely determined by f. If f is a definite form with exactly k automorphs, we shall show that the number of representations of m by f is exactly kN, where N is the number of ideals in $\mathfrak{C}$ of norm m. We then show that if m is prime to $2\alpha\Delta$, the number of ideals in $\mathfrak{G}$ of norm m is exactly $\sigma(m)$, the sum of the positive divisors of m.

Let $f_1, f_2, \cdots, f_H$ be representatives of those classes of forms mentioned above and assume that they are definite forms. If m is positive and prime to $2\alpha\Delta$, if k_i is the number of automorphs of f_i $(i=1, 2, \cdots, H)$ and if every representation of m by f_i be counted as $1/k_i$, then, by the above mentioned results, the total number of representations of m by all the f's is exactly $\sigma(m)$.

We also prove a theorem, which includes one due to Humbert, on certain "generalized" or symbolic representations of an integer by the above f's. It will be seen that such representations may be interpreted explicitly in terms of the corresponding ideals in $\mathfrak{G}$.

There are many results in the literature on the representation of integers by special quaternary quadratic forms which may be written as binary Hermitian forms.‡ A large number of these may be readily obtained from the results of Tr. and of this paper. Several examples are given in the last paragraph.

---

* Presented to the Society, April 10, 1936; received by the editors March 25, 1936.

† *On ideals in generalized quaternion algebras and Hermitian forms*, these Transactions, vol. 38 (1935), pp. 436–446. This paper will be referred to hereafter as Tr.

‡ Cf. Dickson, *History of the Theory of Numbers*, vol. 3, Chapter 10.

2. **The representation of an integer by an Hermitian form.** We shall use the same definitions and notations as used in Tr. unless the contrary is explicitly stated. It will be assumed, as in Tr., that all ideals referred to are left ideals. We shall also assume throughout the paper that all ideals considered are regular ideals.

Let

$$f(x, y) \equiv axx' + bx'y + b'xy' + cyy'$$

be an Hermitian form over G, which represents positive integers, of determinant $bb' - ac = \alpha$, a and c being rational integers and b a number in G. Let m be a positive integer. If $f(r, s) = m$, where r and s are in G, f will be said to represent m. We shall consider the representations of m by f.

Let $f'(x, y) \equiv f(x', y')$. By Theorem 3 of Tr., there is a uniquely determined class $\mathcal{C}$ of ideals in $\mathfrak{G}$ corresponding to the class of forms containing f'. By §5 of Tr., we have

LEMMA 1. *An ideal of $\mathfrak{G}$ is in $\mathcal{C}$ if and only if it has a proper basis Z_1, Z_2 such that*

$$\begin{aligned} EZ_1 &= -b'Z_1 + aZ_2, \\ EZ_2 &= -cZ_1 + bZ_2. \end{aligned} \tag{1}$$

If Z_1, Z_2 form a proper basis of an ideal $\mathfrak{L}$, then the Z's satisfy (1) *if and only if* $N(xZ_1 + yZ_2) = N(\mathfrak{L})f'(x, y)$.*

We shall next prove

LEMMA 2. *If r, s are numbers in G such that $f(r, s) = m > 0$, then*

$$\begin{aligned} \omega_1 &\equiv ar + bs - Es = ar + bs - s'E, \\ \omega_2 &\equiv b'r + cs + Er = b'r + cs + r'E, \end{aligned} \tag{2}$$

form a proper basis of an ideal $\mathfrak{L}$ in $\mathcal{C}$ and $N(\mathfrak{L}) = m$. If $\mathfrak{L}$ is in $\mathcal{C}$ and $N(\mathfrak{L}) = m$, then $\mathfrak{L}$ has a proper basis ω_1, ω_2 as above, where r and s are in G and $f(r, s) = m$.

Suppose $f(r, s) = m > 0$. It may be verified that ω_1, ω_2 in (2) satisfy (1). Hence they form a basis of an ideal $\mathfrak{L}$ in $\mathfrak{G}$. Since the determinant

$$\begin{vmatrix} ar + bs & -s' \\ b'r + cs & r' \end{vmatrix} = f(r, s) = m,$$

it follows that the ω's form a proper basis and $N(\mathfrak{L}) = m$. By Lemma 1, $\mathfrak{L}$ is in $\mathcal{C}$. We shall say that the representation $f(r, s) = m$ belongs to $\mathfrak{L}$.

* The statement "there is a one-to-one correspondence · · · " in lines 8, 9, p. 442, Tr. is false. There is a correspondence but it is not one-to-one. For if the transformation (8) of Tr. is an automorph of f, then f corresponds to the bases ω_1, ω_2 and ξ_1, ξ_2. This error does not affect the validity of any subsequent proof or theorem.

Suppose $\mathfrak{L}$ is an ideal in $\mathcal{C}$ of norm m. By Lemma 1, $\mathfrak{L}$ has a proper basis $\omega_i = g_{i1} + g_{i2}E$ $(i = 1, 2)$, where the ω's satisfy (1) and $N(x\omega_1 + y\omega_2) = mf'(x, y)$. We have $m = N(\mathfrak{L}) = |g_{ij}| = g_{22}\omega_1 - g_{12}\omega_2$. Hence $f'(g_{22}, -g_{12}) = f(g_{22}', -g_{12}') = m$. Replace each Z_i in (1) by $g_{i1} + g_{i2}E$ and equate the coefficients of E. We obtain the first pair of equations below.

$$g_{11} = ag_{22}' - bg_{12}', \qquad \omega_1 = ag_{22}' - bg_{12}' + g_{12}E,$$
$$g_{21} = b'g_{22}' - cg_{12}', \qquad \omega_2 = b'g_{22}' - cg_{12}' + g_{22}E.$$

Noting that $f(g_{22}', -g_{12}') = m$ and comparing the last pair of equations with (2), the lemma follows.

Let $\mathfrak{L}$ be an ideal in $\mathcal{C}$ of norm m and let $f(r, s) = m$ be an arbitrarily chosen but fixed representation belonging to $\mathfrak{L}$. We shall consider the set of all representations of m by f which belong to $\mathfrak{L}$.

Let T be an automorph of f, i.e., a transformation

$$\begin{aligned} x &= t_{11}x_1 + t_{21}y_1, \\ y &= t_{12}x_1 + t_{22}y_1 \end{aligned} \tag{3}$$

which transforms f into itself, the t's being in G and the determinant $|t_{ij}| = 1$. Let $\bar{r}, \bar{s}$ be defined by the equations obtained by replacing x, y, x_1, y_1 in (3) by $r, s, \bar{r}, \bar{s}$ respectively. Then $f(\bar{r}, \bar{s}) = m$. This representation will be said to correspond to T.

We shall show that a representation of m by f belongs to $\mathfrak{L}$ if and only if it corresponds to an automorph of f; also that the representations corresponding to distinct automorphs are distinct. We shall thus have a one-to-one correspondence between those representations of m by f which belong to $\mathfrak{L}$ and the automorphs of f.

LEMMA 3. *A representation of m by f belongs to $\mathfrak{L}$ if and only if it corresponds to an automorph.*

Consider the representation $f(\bar{r}, \bar{s}) = m$ which corresponds to T. Then, employing (2), we have

$$\begin{aligned} \bar{\omega}_1 &\equiv t_{11}'\omega_1 + t_{12}'\omega_2 = g_{11} - E\bar{s}, \\ \bar{\omega}_2 &\equiv t_{21}'\omega_1 + t_{22}'\omega_2 = g_{21} + E\bar{r}, \end{aligned} \tag{4}$$

where g_{11}, g_{21} are certain numbers in G. Since $|t_{ij}'| = 1$, the $\bar{\omega}$'s form a proper basis of $\mathfrak{L}$ by §4 of Tr. We have $x_1\bar{\omega}_1 + y_1\bar{\omega}_2 = x\omega_1 + y\omega_2$, where

$$\begin{aligned} x &= t_{11}'x_1 + t_{21}'y_1, \\ y &= t_{12}'x_1 + t_{22}'y_1. \end{aligned} \tag{5}$$

Since (3) is an automorph of f, (5) is an automorph of f'. Hence $N(x_1\bar{\omega}_1+y_1\bar{\omega}_2) = mf'(x_1, y_1)$. By Lemma 1, the $\bar{\omega}$'s satisfy (1). Therefore, as in the second part of the proof of Lemma 2, we may express g_{11}, g_{21} in terms of $\bar{r}$, $\bar{s}$ and find

$$\begin{aligned} \bar{\omega}_1 &= a\bar{r} + b\bar{s} - E\bar{s}, \\ \bar{\omega}_2 &= b'\bar{r} + c\bar{s} + E\bar{r}. \end{aligned} \tag{6}$$

Comparing these equations with (2), we see that the representation $f(\bar{r}, \bar{s}) = m$ belongs to $\mathfrak{L}$.

Conversely, suppose $f(\bar{r}, \bar{s}) = m$ is a representation which belongs to $\mathfrak{L}$. The $\bar{\omega}$'s of (6) satisfy (1) and form a proper basis of $\mathfrak{L}$. Then they may be expressed in terms of the ω's as in (4), the t_{ij}' being in G and $|t_{ij}'| = 1$. It may be shown that (5) is an automorph of f' and hence (3) is an automorph of f. If in the first pair of equations (4) we replace each ω_i, $\bar{\omega}_j$ by the expressions (2) and (6) and equate the coefficients of E, we see that $f(\bar{r}, \bar{s}) = m$ is the representation which corresponds to the automorph (3). This completes the proof of the lemma.

Suppose (k_{ji}) is the matrix of an automorph of f. Let $\bar{\omega}_i \equiv k_{i1}'\omega_1 + k_{i2}'\omega_2$ $(i = 1, 2)$, where the ω's are given by (2). Then, as in the preceding proof, the $\bar{\omega}$'s satisfy (1) and we have

$$\begin{aligned} E\bar{\omega}_1 &= -b'\bar{\omega}_1 + a\bar{\omega}_2, & E\omega_1 &= -b'\omega_1 + a\omega_2, \\ E\bar{\omega}_2 &= -c\bar{\omega}_1 + b\bar{\omega}_2, & E\omega_2 &= -c\omega_1 + b\omega_2. \end{aligned}$$

In the first pair of equations, set each $\bar{\omega}_i = k_{i1}'\omega_1 + k_{i2}'\omega_2$, replace each $Ek_{ij}'\omega_j$ by $k_{ij}E\omega_j$, and eliminate the $E\omega_j$ by use of the second pair of equations. Employing the left linear independence of the ω's with respect to G (§4, Tr.), we obtain a system of equations equivalent to the following:*

$$\begin{aligned} bk_{11}' + ck_{12}' &= bk_{11} - ak_{21}, & bk_{21}' + ck_{22}' &= ck_{11} - b'k_{21}, \\ ak_{11}' + b'k_{12}' &= -bk_{12} + ak_{22}, & ak_{21}' + b'k_{22}' &= -ck_{12} + b'k_{22}. \end{aligned} \tag{7}$$

We shall now prove

LEMMA 4. *Let $f(\bar{r}, \bar{s}) = m$, $f(R, S) = m$ be the representations belonging to $\mathfrak{L}$ which correspond to the distinct automorphs T, T_1 respectively. Then $(\bar{r}, \bar{s}) \neq (R, S)$.*

We have

$$\begin{aligned} \bar{r} &= k_{11}R + k_{21}S, \\ \bar{s} &= k_{12}R + k_{22}S, \end{aligned} \tag{8}$$

where (k_{ji}) is the matrix of the automorph $T^{-1}T_1$. Suppose $(\bar{r}, \bar{s}) = (R, S)$.

* The condition $|k_{ij}| = 1$ and (7) imply that (k_{ij}) is the matrix of an automorph of f. If $a \neq 0$, $|k_{ij}| = 1$ and the first pair of (7) imply the second pair.

By (8), $(k_{11}-1)(k_{22}-1)-k_{12}k_{21}=0$. Since $|k_{ji}|=1$, $k_{11}+k_{22}=2$. By (8_2), $k_{12}R=(k_{11}-1)S$. We have $k_{12}k_{12}'m=f(k_{12}R,\ k_{12}S)=f(k_{11}-1,\ k_{12})SS'$. But $f(k_{11}-1,\ k_{12})=f(k_{11},\ k_{12})-(k_{11}+k_{11}')a-(k_{12}b+k_{12}'b')+a$. Since $T^{-1}T_1$ is an automorph, $f(k_{11},\ k_{12})=a$. By (7_2), $k_{12}b+k_{12}'b'=(k_{22}-k_{11}')a$. Therefore $f(k_{11}-1, k_{12})=2a-(k_{11}+k_{22})a=0$, $k_{12}=0$ and $k_{11}=k_{22}=1$. From $k_{21}S=(k_{22}-1)R$ we find similarly $k_{21}=0$. Therefore $T^{-1}T_1$ is the identity transformation and $T=T_1$, contrary to hypothesis. The lemma follows.

By Lemmas 2, 3, and 4, we have

THEOREM 1. *Let $f(x, y)$ be an Hermitian form over G, of determinant α, which represents positive integers, let $f'(x, y)\equiv f(x', y')$, and let $\mathcal{C}$ be the class of ideals in $\mathfrak{G}$ which corresponds to the class of forms containing f'. Every representation of a positive integer m by f belongs to a uniquely determined ideal $\mathfrak{L}$ in $\mathcal{C}$ of norm m and for every such ideal $\mathfrak{L}$, there is a representation of m by f which belongs to $\mathfrak{L}$. Those representations of m by f which belong to the same ideal may be placed in one-to-one correspondence with the automorphs of f.*

If f is a definite form, it has only a finite number of automorphs and we have the

COROLLARY. *If f is a definite form, if k is the number of automorphs of f and if $N(\geqq 0)$ is the number of ideals in $\mathcal{C}$ of norm m, then the number of representations of m by f is exactly kN.*

3. The number of ideals with given components. If $p_1,\ p_2, \cdots,\ p_n$ are the distinct prime factors of an integer a prime to 2Δ, let

$$\Psi(a) \equiv a\prod_{i=1}^{n}\left[1-\left(\frac{\Delta}{p_i}\right)\frac{1}{p_i}\right],$$

where (Δ/p_i) is Legendre's symbol. Let $\Psi(1)\equiv 1$. If A, B are relatively prime, $\Psi(A)\cdot\Psi(B)=\Psi(AB)$. The following is an immediate consequence of a result due to Hermite.*

LEMMA 5. *If a is an integer prime to $2\alpha\Delta$, there are exactly $\Psi(a)$ numbers X, in G, no two of which are congruent modulo a, such that*

$$N(X)-\alpha\equiv 0 \pmod{a}. \tag{9}$$

By Lemma 2 of Tr., the first and second components of an ideal in $\mathfrak{G}$ are $a\mathfrak{b}$ and $\mathfrak{b}$, where a is a positive rational integer and $\mathfrak{b}$ is an ideal in G. Let a be a positive rational integer prime to $2\alpha\Delta$, and let $\mathfrak{b}$ be an ideal in G. We shall show that there are exactly $\Psi(a)$ ideals in $\mathfrak{G}$ whose components are $a\mathfrak{b}$ and $\mathfrak{b}$.

* *Oeuvres*, vol. 1, pp. 247–250.

Suppose $\mathfrak{L}$ is an ideal in $\mathfrak{G}$ with those components. If $\mathfrak{b}_1$ is equivalent to $\mathfrak{b}$, as in the proof of Lemma 2 of Tr., $\mathfrak{L}t=\mathfrak{L}_1t_1$, where t, t_1 are in G and $\mathfrak{L}_1$ is a uniquely determined ideal in $\mathfrak{G}$ with the components $a\mathfrak{b}_1$ and $\mathfrak{b}_1$. Conversely, if $\mathfrak{L}_1$ is such an ideal, the equation $\mathfrak{L}t=\mathfrak{L}_1t_1$ defines uniquely an ideal $\mathfrak{L}$ with the components $a\mathfrak{b}$ and $\mathfrak{b}$. Since every class of ideals in G contains an ideal prime to an arbitrarily chosen integer, in our proof of the above-mentioned result, we may assume without loss of generality that $\mathfrak{b}$ is prime to $2a\alpha\Delta$.

Let λ_1, λ_2 be a canonical basis of $\mathfrak{b}$. λ_1 is a divisor of $N(\mathfrak{b})$ and hence is prime to $2a\alpha\Delta$. After adding a proper multiple of λ_1 to λ_2, we may assume that λ_2 is also prime to $2a\alpha\Delta$.

Let $\mathfrak{L}$ be an ideal with the components $a\mathfrak{b}$ and $\mathfrak{b}$. By Lemma 1, Tr.,

$$\mathfrak{L} = [a\lambda_1, a\lambda_2, b_1 + E\lambda_1, b_2 + E\lambda_2],$$

where the b's are in $\mathfrak{b}$. If $B_i \equiv b_i \pmod{a\mathfrak{b}}$ $(i=1, 2)$, each b_i above may be replaced by B_i. Consider the congruences

$$\lambda_i \bar{b}_i \equiv b_i \pmod{a\mathfrak{b}} \qquad (i = 1, 2).$$

Since λ_i is prime to a and $\mathfrak{b}$ divides λ_i and b_i, these congruences have solutions which are uniquely determined modulo a. Then we may replace each b_i in the above basis of $\mathfrak{L}$ by $\lambda_i\bar{b}_i$. Dropping the bars and setting $\Omega_i \equiv (b_i+E)\lambda_i$ $(i=1, 2)$, we have

$$\mathfrak{L} = [a\lambda_1, a\lambda_2, \Omega_1, \Omega_2]. \tag{10}$$

It will be understood hereafter that the b_i are as in the definitions of the Ω's and not as in the first basis of $\mathfrak{L}$ given above.

Since each $E\Omega_i$ is in $\mathfrak{L}$ and

$$E\Omega_i = b_i'\Omega_i + [\alpha - N(b_i)]\lambda_i \qquad (i = 1, 2), \tag{11}$$

it follows that

$$N(b_i) - \alpha \equiv 0 \pmod{a} \qquad (i = 1, 2); \tag{12}$$

i.e., each b_i is a solution of (9).

Let G have the basis $(1, \theta)$. Each $\theta\lambda_i$, $\theta'\lambda_i$ is in $\mathfrak{b}$ and therefore

$$\begin{aligned} \theta\lambda_i &= u_{i1}\lambda_1 + u_{i2}\lambda_2, \\ \theta'\lambda_i &= v_{i1}\lambda_1 + v_{i2}\lambda_2, \end{aligned} \qquad (i = 1, 2), \tag{13}$$

where the u's and v's are rational integers. It may be shown that $u_{12}=-v_{12}=\pm N(\lambda_1)/N(\mathfrak{b})$, $u_{21}=-v_{21}=\pm N(\lambda_2)/N(\mathfrak{b})$, $(u_{11}-\theta)(u_{22}-\theta)=(v_{11}-\theta)(v_{22}-\theta)=u_{12}u_{21}=v_{12}v_{21}$. Then $v_{12}v_{21}$ is prime to a. Employing the expressions (13) for $\theta'\lambda_i$ we find

$$\begin{aligned}\theta\Omega_1 &= v_{11}\Omega_1 + v_{12}\Omega_2 + [(\theta - v_{11})\lambda_1 b_1 - v_{12}\lambda_2 b_2],\\ \theta\Omega_2 &= v_{21}\Omega_1 + v_{22}\Omega_2 + [(\theta - v_{22})\lambda_2 b_2 - v_{21}\lambda_1 b_1].\end{aligned} \tag{14}$$

Since the Ω's are in $\mathfrak{L}$ and each of the expressions in brackets above is in G, it follows that

$$\begin{aligned}(\theta - v_{11})\lambda_1 b_1 - v_{12}\lambda_2 b_2 &\equiv 0,\\ v_{21}\lambda_1 b_1 - (\theta - v_{22})\lambda_2 b_2 &\equiv 0,\end{aligned} \qquad (\text{mod } a). \tag{15}$$

Since $(\theta - v_{11})(\theta - v_{22}) - v_{12}v_{21} = 0$ and $v_{12}v_{21}$ is prime to a, it follows that the congruences (15) are equivalent. From (15_1) and (13) respectively we have

$$\begin{aligned}N(\theta - v_{11})N(\lambda_1)N(b_1) &\equiv v_{12}^2 N(\lambda_2)N(b_2) \qquad (\text{mod } a)\\ N(\theta - v_{11})N(\lambda_1) &= v_{12}^2 N(\lambda_2).\end{aligned}$$

Hence (15_1) implies $N(b_1) \equiv N(b_2) \pmod{a}$. It follows that (12_1) and (15_1) imply (12_2) and (15_2).

Conversely, let b_1 be a root of (9). Since $v_{12}\lambda_2$ is prime to a, b_2 is uniquely determined, modulo a, by (15_1) and b_1, b_2 satisfy each of the four congruences (12) and (15). Let $\Omega_i \equiv (b_i + E)\lambda_i$ $(i = 1, 2)$. Since a is prime to $\mathfrak{b}$, it follows from (11) and (14) that each $E\Omega_i$, $\theta\Omega_i$ is a linear function, with rational integral coefficients, of $a\lambda_1$, $a\lambda_2$, Ω_1, Ω_2. The same is obviously true of each $Ea\lambda_i$, $\theta a\lambda_i$. Hence there is an ideal $\mathfrak{L}$ with a basis as in (10) with $a\mathfrak{b}$ and $\mathfrak{b}$ as its components. Let b_1 be another root of (9). We may obtain in a similar manner an ideal $\mathfrak{L}_1$. It may be shown that $\mathfrak{L}_1 = \mathfrak{L}$ if and only if $b \equiv b_1 \pmod{a}$. We have then by Lemma 5

THEOREM 2. *If a is a positive integer, prime to $2\alpha\Delta$, and $\mathfrak{b}$ is an ideal in G, there are exactly $\Psi(a)$ ideals in $\mathfrak{G}$ whose first and second components are $a\mathfrak{b}$ and $\mathfrak{b}$ respectively.*

4. **The number of ideals with a given norm.** We shall prove

THEOREM 3. *If m is a positive integer, prime to $2\alpha\Delta$, there are exactly $\sigma(m)$ ideals in $\mathfrak{G}$ of norm m.*

Since the theorem is obviously true for $m = 1$, it will be sufficient to show that if it is true for $m = m_1$, then it is true for $m = m_1 P^t$, where P is a prime not a divisor of m_1.

By the proof of Lemma 3 of Tr., for every ideal $\mathfrak{L}$ in $\mathfrak{G}$, there is a ρ in $\mathfrak{G}$ of positive norm such that $\mathfrak{L}\rho = \mathfrak{L}_1$, where $\mathfrak{L}_1$ is a reduced ideal and $n(\mathfrak{L})N^2(\rho) = n(\mathfrak{L}_1)$. Furthermore, by p. 440 of Tr., $N(\mathfrak{L})N(\rho) = N(\mathfrak{L}_1)$, $\mathfrak{L}_1 = [a_1, b_1 + E]$ and $N(\mathfrak{L}_1) = a_1$. Since $\mathfrak{L}_1 = [a_1, a_1\theta, b_1 + E, \theta(b_1 + E)]$, by Lemma 1 of Tr., $n(\mathfrak{L}_1) = a_1^2$. By the same reference, $n(\mathfrak{L}) = a^2 N^2(\mathfrak{b})$, where $a\mathfrak{b}$ and $\mathfrak{b}$ are the components of $\mathfrak{L}$. Therefore $N(\mathfrak{L}) = aN(\mathfrak{b})$. Hence by Theo-

rem 2, the number $\phi(m)$, of ideals in $\mathfrak{G}$ of norm m is $\sum\Psi[m/N(\mathfrak{b})]$ where the summation extends over all the ideals $\mathfrak{b}$ in G whose norms divide m.

Suppose $\phi(m_1)=\sigma(m_1)$. Let a be a positive integer and $\mathfrak{b}$ an ideal in G such that $aN(\mathfrak{b})=m_1$. Let P be a rational prime not a divisor of $2\alpha\Delta m_1$, and let t be a positive integer.

Suppose $(\Delta/P)=1$. For every integer $u\geqq 0$, there are exactly $u+1$ ideals in G of norm P^u. Let $0\leqq u\leqq t$ and let $\mathfrak{p}$ be an ideal of norm P^u. By Theorem 2, there are exactly $\Psi(aP^{t-u})=\Psi(a)\cdot\Psi(P^{t-u})$ ideals in $\mathfrak{G}$ with the components $aP^{t-u}\mathfrak{b}\mathfrak{p}$ and $\mathfrak{b}\mathfrak{p}$ and each of these ideals is of norm m_1P^t. But $\mathfrak{p}$ may be chosen in $u+1$ ways. Hence there are exactly

$$\Psi(a)\sum_{u=0}^{t}(u+1)\Psi(P^{t-u})=\Psi(a)\sigma(P^t)$$

ideals in $\mathfrak{G}$ of norm m_1P^t whose second components are $\mathfrak{b}\mathfrak{p}$, where $\mathfrak{p}$ is an ideal whose norm is a power ($\geqq 0$) of P. By Theorem 2, there are exactly $\Psi(a)$ ideals of norm m_1 whose second components are $\mathfrak{b}$. And by hypothesis $\phi(m_1)=\sigma(m_1)$. It follows that

$$\phi(m_1P^t)=\sigma(m_1)\sigma(P^t)=\sigma(m_1P^t).$$

Suppose $(\Delta/P)=-1$. Then $\{P\}$ is a prime in G. If $0\leqq v\leqq t/2$, by Theorem 2 there are exactly $\Psi(aP^{t-2v})=\Psi(a)\cdot\Psi(P^{t-2v})$ ideals in $\mathfrak{G}$ of norm m_1P^t whose second components are $\mathfrak{b}P^v$. Since

$$\Psi(a)\sum_{v=0}^{[t/2]}\Psi(P^{t-2v})=\Psi(a)\sigma(P^t),$$

where $[t/2]$ is the greatest integer $\leqq t/2$, it follows as above that

$$\phi(m_1P^t)=\sigma(m_1)\sigma(P^t)=\sigma(m_1P^t)$$

and the theorem is proved.

4. **The representation of an integer by a system of forms.** By Theorem 3 of Tr., there is a one-to-one correspondence between the classes of ideals in $\mathfrak{G}$ and those classes of binary Hermitian forms over G, of determinant α, which represent positive integers. Let $f_1, f_2, \cdots, f_H$ be representatives of those H classes of such forms which represent integers prime to $2\alpha\Delta$.

It may be shown that no two of the forms $f_i'(x,y)\equiv f_i(x',y')$ $(i=1,2,\cdots,H)$ are equivalent. Since each f_i' represents the same integers represented by f_i, it follows that they also form a system of representative forms of the above H classes. Let $\mathcal{C}_1, \mathcal{C}_2, \cdots, \mathcal{C}_H$ be the classes of ideals in $\mathfrak{G}$ corresponding to the classes containing $f_1', f_2', \cdots, f_H'$ respectively. Let m be a positive integer prime to $2\alpha\Delta$. If $\mathfrak{L}$ is an ideal of norm m in the class $\mathcal{C}$, every form in the

corresponding class represents m. Hence $\mathcal{C}$ is one of the $\mathcal{C}_i$. Assume that the f_i are definite forms, i.e., $\alpha<0, \Delta<0$. By the corollary of Theorem 1, the number of representations of m by a given f_i is k_iN_i, where k_i is the number of automorphs of f_i and N_i is the number of ideals in $\mathcal{C}_i$ of norm m. By Theorem 3, $N_1+N_2+\cdots+N_H=\sigma(m)$. Hence, if every representation of m by f_i be counted as $1/k_i$, the total number of representations of m by $f_1, f_2, \cdots, f_H$ is exactly $\sigma(m)$.

Consider those representations $f_i(r, s)=m$ by a given f_i in which r, s are relatively prime. Such a representation is said to be proper. By Lemma 2 and (2), every such representation belongs to an ideal $\mathfrak{L}$ of norm m in $\mathcal{C}_i$ whose second component is the unit ideal. Then the first component of $\mathfrak{L}$ is $\{m\}$. Conversely, by the same reference, for every such ideal $\mathfrak{L}$ in $\mathcal{C}_i$ there is a proper representation of m by f_i which belongs to $\mathfrak{L}$. By Theorem 2, there are exactly $\Psi(m)$ such ideals $\mathfrak{L}$ in $\mathfrak{G}$. We have then

THEOREM 4. *Let $f_1, f_2, \cdots, f_H$ be representatives of the H classes of binary Hermitian forms over G which represent positive integers prime to $2\alpha\Delta$, and assume that $\alpha<0, \Delta<0$. Let k_i be the number of automorphs of f_i $(i=1, 2, \cdots, H)$ and let m be a positive integer prime to $2\alpha\Delta$. The total number of [proper] representations of m by $f_1, f_2, \cdots, f_H$ is exactly $\sigma(m)$ $[\Psi(m)]$ if a representation by f_i be counted as $1/k_i$.**

5. **An extension of a theorem of Humbert's.** Let f be one of the f_i in Theorem 4 and let $\mathfrak{p}$ be an ideal in G. If r and s, not both zero, are in $\mathfrak{p}$ then $f(r, s)=mN(\mathfrak{p})$ where m is a positive integer. Humbert wrote symbolically $m=f(r/\mathfrak{p}, s/\mathfrak{p})$ and called this a generalized representation of m belonging to $\mathfrak{p}$. If $\mathfrak{p}$ is the g.c.d. of r, s, this representation is said to be proper. He gave the following

THEOREM. (Humbert) *If Δ is even, if m is positive and prime to $2\alpha\Delta$, and if $\mathfrak{p}_1, \mathfrak{p}_2, \cdots, \mathfrak{p}_h$ are representatives of the h classes of ideals in G, then the total number of generalized representations of m belonging to $\mathfrak{p}_1, \mathfrak{p}_2, \cdots, \mathfrak{p}_h$ by $f_1, f_2, \cdots, f_H$ is exactly $h\sigma(m)$, where a representation by f_i is counted as $1/k_i$.†*

We shall obtain a stronger theorem. Let $\mathfrak{p}$ be an ideal in G and let m be a positive integer prime to $2\alpha\Delta$. Suppose $a\mathfrak{b}$ and $\mathfrak{b}$ are the components of an ideal in $\mathfrak{G}$ of norm m. By Theorem 2, there are exactly $\Psi(a)$ such ideals in $\mathfrak{G}$. But by the same reference, there are exactly $\Psi(a)$ ideals in $\mathfrak{G}$ whose com-

* For the case where Δ is even, Humbert obtained the above result on the proper representations of m. See Comptes Rendus, Paris, vol. 169 (1919), pp. 309–315.

† Comptes Rendus, Paris, vol. 169 (1919), p. 365. We state the theorem in our notation. Our Δ, α, $\mathfrak{p}_i$ are Humbert's $-4P$, Δ, and I_i respectively. Humbert does not state explicitly that Δ is even but this is implied by the former condition that $P\equiv 1$ or 2 (mod 4).

ponents are $a\mathfrak{b}\mathfrak{p}$ and $\mathfrak{b}\mathfrak{p}$, and the norm of every such ideal is $mN(\mathfrak{p})$. It follows from Theorem 3 that there are exactly $\sigma(m)$ ideals in $\mathfrak{G}$ of norm $mN(\mathfrak{p})$ whose second components are divisible by $\mathfrak{p}$. By Theorem 2, $\mathfrak{p}$ is the second component of exactly $\Psi(m)$ of these ideals.

Let f be one of the f_i of Theorem 4 and let $\mathcal{C}$ be the class of ideals in $\mathfrak{G}$ determined by f'. By Lemma 2 and (2), every representation $f(r, s)=mN(\mathfrak{p})$, where r, s are in $\mathfrak{p}$ belongs to an ideal $\mathfrak{L}$ in $\mathcal{C}$ of norm $mN(\mathfrak{p})$ whose second component is divisible by $\mathfrak{p}$. Then $f(r/\mathfrak{p}, s/\mathfrak{p})=m$ is a generalized representation of m belonging to $\mathfrak{p}$ and this representation is proper if the second component of $\mathfrak{L}$ is $\mathfrak{p}$. Conversely, by the same reference, for every such ideal $\mathfrak{L}$, there is a representation $f(r, s)=mN(\mathfrak{p})$ which belongs to $\mathfrak{L}$ and the generalized representation $f(r/\mathfrak{p}, s/\mathfrak{p})=m$ is proper only if $\mathfrak{p}$ is the second component of $\mathfrak{L}$. We have then

THEOREM 5. *Let m be a positive integer prime to $2\alpha\Delta$, let $\mathfrak{p}$ be an ideal in G, and let f_i, k_i be as in Theorem* 4. *The total number of generalized* [*proper*] *representations of m belonging to $\mathfrak{p}$ by $f_1, f_2, \cdots, f_H$ is exactly $\sigma(m)$* [$\Psi(m)$], *it being understood that a representation by f_i is counted as $1/k_i$.*

It will be observed that for the case $\mathfrak{p}=\{1\}$, Theorem 5 becomes Theorem 4; also, that Theorem 5 includes Humbert's theorem quoted above.

6. **Applications.** Consider the case $\Delta=-4$, $\alpha=-1$. By p. 445 of Tr., $H=1$ and we may take $f_1=xx'+yy'$. f_1 has exactly eight automorphs corresponding to the eight units in $\mathfrak{G}$. Hence we have by Theorem 4, Jacobi's well known result that the number of representations of an odd positive integer m by $u^2+v^2+z^2+w^2$ is exactly $8\sigma(m)$. We also note that by Theorem 4, the number of representations with z^2+w^2 prime to m is $\Psi(m)$. Similarly, for the case $\Delta=-8$, $\alpha=-1$, we obtain Liouville's result that the number of representations of m by $u^2+v^2+2z^2+2w^2$ is $4\sigma(m)$.*

Consider the case $\Delta=-7$, $\alpha=-1$. By p. 445 of Tr., all the regular ideals in $\mathfrak{G}$ are principal and by Lemma 6 of Tr., all the ideals in $\mathfrak{G}$ are regular. Then $H=1$, $f_1=xx'+yy'$. There are exactly three ideals in $\mathfrak{G}$ of norm 2, namely, those defined by $1+E$, $\theta=(1+\Delta^{1/2})/2$ and θ'. More generally, it may be shown by induction that the number of ideals of norm 2^n is $2^{n+1}-1=\sigma(2^n)$. Furthermore, there is only one ideal of norm 7^n, namely, that defined by $(2\theta-1)^n$. Then, employing Theorem 3 of this paper and Theorem 4 of Tr., it may be shown that if $m=7^n m_1$, where m_1 is a positive integer prime to 7, the number of ideals of norm m is exactly $\sigma(m_1)$. Since $\mathfrak{G}$ contains exactly four units, it follows that the number of representations of m by f_1 is exactly $4\sigma(m_1)$. Noting that f_1 may be written $f_1=u^2+uz+2z^2+v^2$

* Journal de Mathématiques, (2), vol. 5 (1860), p. 270.

$+vw+2w^2$, we see that this result was obtained by Dickson.* If m is prime to $2\cdot 7$, by Theorem 4 there are exactly $4\Psi(m)$ representations with $v^2+vw+2w^2$ prime to m.

Consider the case $\Delta=-8$, $\alpha=-3$. Let m be a positive integer prime to 6. It may be shown that $H=k_1=k_2=2$ and that we may take $f_1=xx'+3yy'$, $f_2=2xx'+(1+\theta)x'y+(1-\theta)xy'+3yy'$, where $\theta^2=-2$. By Theorem 4, the total number of representations of m by f_1 and f_2 is exactly $2\sigma(m)$.

f_1 may be written $u^2+2v^2+3z^2+6w^2$. The problem of determining the number $T(m)$, of representations of m by f_1 has long been recognized as a very difficult problem. Liouville stated that $T(m)\leqq 2\sigma(m)$,† an immediate consequence of our result above. By the corollary of Theorem 1, the problem of determining $T(m)$ is equivalent to that of determining how many of the $\sigma(m)$ ideals of norm m are principal. This is the same sort of problem which arises in the determination of the integers represented by a binary quadratic form when there is more than one class in a genus.

By a result due to Griffiths,‡ $T(m)<2\sigma(m)$ if $m>1$. Since $f_2(\theta x, \theta y)=2f_2(x, y)$, $f_2[(1+\theta)x, (1+\theta)y]=3f_2(x, y)$, it follows that f_2 represents every integer >1.

* *Algebren und ihre Zahlentheorie*, p. 197.

† Journal de Mathématiques, (2), vol. 9 (1864), p. 305.

‡ American Journal of Mathematics, vol. 51 (1929), p. 66, Theorem 6.

UNIVERSITY OF KENTUCKY,
LEXINGTON, KY.

ON SOME HERMITIAN FORMS ASSOCIATED WITH TWO GIVEN CURVES OF THE COMPLEX PLANE*

BY

GABRIEL SZEGÖ

1. Let us consider the set of polynomials $f(z)$ of degree n satisfying the condition $|f(z)| \leqq 1$ on a given closed set E. What can then be said about $|f(z)|$ on another given closed set E'?

The maximum of $|f(z)|$ on E' for the set of all polynomials defined by the inequality mentioned is a number $m_n = m_n(E, E')$ which depends on n and also on the mutual position of E and E'. It is of course an invariant with respect to translation and rotation. Obviously $m_n = 1$ if E' is a subset of E. For the sake of simplicity we further assume that E and E' are closed regions bounded by rectifiable Jordan curves C and C'. In case C' has some points exterior to C, the exact determination of m_n seems to be rather difficult. The asymptotic calculation for large values of n, however, is possible. Fekete [1] has shown the existence of

$$\lim_{n\to\infty} (m_n)^{1/n} = \rho. \tag{1}$$

Here the number ρ can be characterized by means of the conformal and one-to-one representation of the region exterior to C onto the region exterior to a circle in the w-plane, $z = \infty$ corresponding to $w = \infty$. The points of C' outside of C have some correspondents in the w-plane. The greatest distance of these from the center of the circle, divided by the radius of the circle, is equal to ρ.

2. An analogous problem can be proposed if the maximum value of $|f(z)|$ on C is replaced by its "mean quadratic value,"

$$\left\{ l^{-1} \int_C |f(z)|^2 |dz| \right\}^{1/2};$$

here C is supposed to have the length l. A similar definition of the deviation on C' may be used. The corresponding maximum $M_n = M_n(C, C')$ depends again on the mutual position of C and C'. On introducing the normalized orthogonal polynomials $\{q_\nu(z)\}$ associated with the curve C [8] it can be easily shown that

* Presented to the Society, April 10, 1936; received by the editors February 24, 1936.

$$M_n^2 = \max \sum_{\mu,\nu=0}^{n} x_\mu \bar{x}_\nu (l')^{-1} \int_{C'} q_\mu(z)\overline{q_\nu(z)} \, |dz|, \tag{2}$$

provided that the variables x_ν satisfy the condition $\sum_{\nu=0}^{n} |x_\nu|^2 \leqq 1$ (or $=1$). In other words, M_n^2 is the greatest characteristic value of the Hermitian form on the right side of (2). Here l' denotes the length of C'.

3. In the present paper we shall not enter into the details of this general problem. [The answer which can be expected is, incidentally, the same as in (1).] Only a special case shall be treated here, in which C is the real interval $-1, +1$ (actually a limiting case of a Jordan curve) and C' the unit circle with the origin as center. We derive then, for the corresponding number M_n the following asymptotic expression:*

$$M_n \cong 2^{-9/8}\pi^{-3/4}n^{-1/4}(2^{1/2}+1)^{n+3/2}. \tag{3}$$

It can be easily shown, that in this case $\rho = 2^{1/2}+1$, ρ having the same meaning as above. Another formulation of this result is: Let λ_n be the smallest characteristic value of the quadratic form

$$\frac{1}{2}\int_{-1}^{1} (x_0 + x_1 l + x_2 l^2 + \cdots + x_n l^n)^2 dl = \sum_{\mu,\nu=0}^{n}{}' \frac{x_\mu x_\nu}{\mu+\nu+1}, \tag{4}$$

where $\sum'$ indicates that the summation is extended only over even values of $\mu+\nu$. We have then

$$\lambda_n \cong 2^{9/4}\pi^{3/2}n^{1/2}(2^{1/2}-1)^{2n+3}, \tag{5}$$

since $\lambda_n = M_n^{-2}$.

In a letter on November 19, 1935, Professor Tamarkin proposed the problem of determining an asymptotic expression for these numbers λ_n. I dealt with this question several years ago in connection with some investigations about the so-called Hankel forms [7]. I found at that time instead of (5), the less sharp result (18) of Part I. The letter of Professor Tamarkin gave me the opportunity to attack this problem again and to obtain the more exact formula (5) as well as the subsequent related results.

The reader may be interested in the facts previously known about the form (4). Hilbert [3] obtained an explicit expression for its determinant

$$D_n = 2^{n(n+1)} \frac{(1!2!\cdots n!)^4}{1!2!3!\cdots(2n+1)!}. \tag{6}$$

Moreover, it is known† that the greatest characteristic value Λ_n tends to $\pi/2$

* In what follows, we write $a_n \cong b_n$ (a_n and b_n real or complex, $\neq 0$) if $\lim_{n\to\infty} a_n b_n^{-1} = 1$; furthermore, $a_n \sim b_n$ (a_n and b_n positive) if two positive constants u and v exist such that $u < a_n b_n^{-1} < v$.

† See, for instance, G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge, 1934, p. 226.

if $n \to \infty$. It would be extremely interesting to find a simple way of describing asymptotically the distribution of all characteristic values of this form for large values of n.

4. If the range $(-1, +1)$ is replaced by another segment, similar results are valid; the asymptotic behavior of the numbers corresponding to λ_n depends, of course, on the situation of this segment. For instance, in case of the form

$$\int_0^1 (x_0 + x_1 t + x_2 t^2 + \cdots + x_n t^n)^2 dt = \sum_{\mu,\nu=0}^{n} \frac{x_\mu x_\nu}{\mu + \nu + 1} \tag{7}$$

we obtain for the smallest characteristic value

$$\lambda_n \cong 2^{15/4} \pi^{3/2} n^{1/2} (2^{1/2} - 1)^{4n+4}. \tag{8}$$

[We have in this case $\rho = (2^{1/2}+1)^2$.] The greatest characteristic value tends here to π.

5. For infinite sets, the integrals considered are not convergent. Thus the introduction of a proper weight factor becomes necessary. The special cases

$$\int_{-\infty}^{\infty} e^{-t^2} (x_0 + x_1 t + x_2 t^2 + \cdots + x_n t^n)^2 dt = \sum_{\mu,\nu=0}^{n}{}' \Gamma(\tfrac{1}{2}(\mu + \nu + 1)) x_\mu x_\nu \tag{9}$$

and

$$\int_0^{\infty} e^{-t} (x_0 + x_1 t + x_2 t^2 + \cdots + x_n t^n)^2 dt = \sum_{\mu,\nu=0}^{n} (\mu + \nu)! x_\mu x_\nu \tag{10}$$

illustrate the totally different behavior of the smallest characteristic value with which we are here concerned. In these cases the corresponding asymptotical values are:*

$$2^{13/4} \pi^{3/2} e n^{1/4} \exp(-2(2n)^{1/2}) \text{ and } 2^{3/2} \pi^{3/2} e n^{1/4} \exp(-4n^{1/2}). \tag{11}$$

6. In Part I we start with the discussion of the form (4) proving that $\lambda_n \sim n^{1/2}(2^{1/2}-1)^{2n}$. We give in Part II the proof of the sharper formula (5). In Part III the analogous treatment of the form (7) follows. Finally, we discuss the forms (9) and (10), giving the proof of (11). The formulas (5) and (8) are connected with asymptotic properties of Legendre polynomials; in cases (9) and (10) we need some asymptotic properties of Hermite and Laguerre polynomials.

PART I

1. Obviously we have in case of the form (4)

* It is easy to show that the corresponding greatest characteristic values are $\cong \Gamma(n+\frac{1}{2})$ and $(2n)!$.

$$\lambda_n = \min \frac{\frac{1}{2}\int_{-1}^{1}(f(t))^2dt}{(1/(2\pi))\int_{-\pi}^{\pi}|f(e^{i\phi})|^2d\phi},$$

or what is the same,

$$\pi\lambda_n^{-1} = \max \frac{\int_{-\pi}^{\pi}|f(e^{i\phi})|^2d\phi}{\int_{-1}^{1}(f(t))^2dt},$$

$f(z)$ being an arbitrary polynomial of the nth degree with real coefficients. Supposing that z lies in the complex plane cut along the segment $(-1, +1)$, we have

$$(12)\quad (z+(z^2-1)^{1/2})^{-2n-1}(f(z))^2 = \frac{1}{2\pi i}\int (t+(t^2-1)^{1/2})^{-2n-1}(f(t))^2\frac{dt}{t-z},$$

where the integral is taken in negative sense along a closed curve enclosing the segment $(-1, +1)$ but not the point z. The left-hand member is indeed regular in the cut plane and vanishes for $z=\infty$. By $z+(z^2-1)^{1/2}$ we mean that branch of the function which is ∞ for $z=\infty$. For this branch $|z+(z^2-1)^{1/2}|>1$ in the cut plane and $=1$ on the cut itself.

Thus we get

$$(13)\qquad |z+(z^2-1)^{1/2}|^{-2n-1}|f(z)|^2 \leqq \frac{1}{\pi\delta}\int_{-1}^{1}(f(t))^2dt,$$

δ being the shortest distance of z from $(-1, +1)$.

Another inequality, less exact in general, can be obtained using some elementary facts about Legendre polynomials. We have [Pólya-Szegö, **5**, vol. 2, p. 95, problem 103]

$$(14)\qquad |(z+(z^2-1)^{1/2})^{-2n}(f(z))^2| \leqq \max_{-1\leqq t\leqq +1}(f(t))^2 \leqq \tfrac{1}{2}(n+1)^2\int_{-1}^{1}(f(t))^2dt.$$

2. The unit circle $|z|=1$ lies entirely interior to the ellipse $|z+(z^2-1)^{1/2}| = 1+2^{1/2}$ having with it only the points $\pm i$ in common. Let γ be an arc $\pi/2-\epsilon\leqq\phi\leqq\pi/2+\epsilon$ of the unit circle containing $+i$ and lying in the upper half plane; let γ' be the complementary arc of the upper half circle. Then we have on account of (13)

$$(15)\qquad \int_{\gamma}|f(z)|^2d\phi \leqq \int_{\gamma}|z+(z^2-1)^{1/2}|^{2n+1}d\phi\cdot\frac{1}{\pi\cos\epsilon}\int_{-1}^{1}(f(t))^2dt.$$

On setting $z=e^{i\phi}=ie^{i\phi'}$ we have in the neighborhood of $z=i$, i.e., for small values of ϕ',

$$\log\left[z+(z^2-1)^{1/2}\right]=\log\left[ie^{i\phi'}+i(e^{2i\phi'}+1)^{1/2}\right]$$

$$=\frac{i\pi}{2}+\log\left[e^{i\phi'}+e^{i\phi'/2}(2\cos\phi')^{1/2}\right]$$

$$=\frac{i\pi}{2}+\log\left\{1+i\phi'-\frac{\phi'^2}{2}+\cdots\right.$$

$$\left.+2^{1/2}\left(1+i\frac{\phi'}{2}-\frac{\phi'^2}{8}+\cdots\right)\left(1-\frac{\phi'^2}{4}+\cdots\right)\right\} \tag{16}$$

$$=\frac{i\pi}{2}+\log(1+2^{1/2})+\log\left\{1+i\frac{\phi'}{2^{1/2}}-(2^{1/2}+2)\frac{\phi'^2}{8}+\cdots\right\}$$

$$=\frac{i\pi}{2}+\log(1+2^{1/2})+2^{-1/2}i\phi'-2^{-1/2}\frac{\phi'^2}{4}+\cdots,$$

whence

$$\left|z+(z^2-1)^{1/2}\right|=\exp\left\{\Re\log\left(z+(z^2-1)^{1/2}\right)\right\}$$

$$=(2^{1/2}+1)\exp\left\{-2^{-1/2}\frac{\phi'^2}{4}+\cdots\right\}. \tag{16'}$$

Using a familiar argument [cf. Pólya-Szegö, **5**, vol. 1, p. 78, problem 201] we obtain

$$\int_\gamma\left|z+(z^2-1)^{1/2}\right|^{2n+1}d\phi\sim n^{-1/2}(2^{1/2}+1)^{2n} \tag{17}$$

and on account of (15)

$$\frac{\int_\gamma|f(z)|^2d\phi}{\int_{-1}^1(f(t))^2dt}=O(n^{-1/2}(2^{1/2}+1)^{2n}).$$

3. On the other hand we obtain on γ' using (14)

$$\frac{|f(z)|^2}{\int_{-1}^1(f(t))^2dt}\leqq\tfrac{1}{2}(n+1)^2\lambda^{2n},$$

λ being the maximum of $|z+(z^2-1)^{1/2}|$ on γ'. But $\lambda<2^{1/2}+1$, whence

$$\frac{\int_{-\pi}^{\pi}|f(e^{i\phi})|^2d\phi}{\int_{-1}^1(f(t))^2dt}=O(n^{-1/2}(2^{1/2}+1)^{2n})$$

follows.

4. For the special polynomial $f(t)=f(\cos\theta)=\cos n\theta$

$$\int_{-1}^1(f(t))^2dt=\int_0^\pi\cos^2n\theta\sin\theta\,d\theta=1-(4n^2-1)^{-1}$$

holds. Furthermore

$$\int_{-\pi}^{\pi} |f(e^{i\phi})|^2 d\phi = \int_{-\pi}^{\pi} \left| \tfrac{1}{2}(z + (z^2-1)^{1/2})^n + \tfrac{1}{2}(z - (z^2-1)^{1/2})^n \right|^2_{z=e^{i\phi}} d\phi.$$

The contribution of γ' to this integral is $=O(\lambda^n)$. On the other hand on γ

$$\frac{1}{4}\int_{\gamma} |z + (z^2-1)^{1/2}|^{2n} |1 + (z - (z^2-1)^{1/2})^{-2n}|^2 d\phi$$

$$> \frac{1}{4}(1-\lambda^{-2n})^2 \int_{\gamma} |z + (z^2-1)^{1/2}|^{2n} d\phi.$$

According to (17) the last expression is $> cn^{-1/2}\,(2^{1/2}+1)^{2n}$, where $c>0$ is independent of n. Thus we have shown

$$\lambda_n \sim n^{1/2}(2^{1/2}-1)^{2n}. \tag{18}$$

PART II

1. In order to prove the more precise statement (5), we introduce the normalized Legendre polynomials $q_n(z) = (n+\frac{1}{2})^{1/2}P_n(z)$, P_n being the ordinary Legendre polynomial of the nth degree. On writing $f(z)$ as a linear combination of Legendre polynomials we obtain

$$\pi\lambda_n^{-1} = \max \sum_{\mu,\nu=0}^{n} K_{\mu\nu} x_\mu \bar{x}_\nu \quad \text{for} \quad \sum_{\nu=0}^{n} |x_\nu|^2 \leqq 1, \tag{19}$$

where

$$K_{\mu\nu} = \int_{-\pi}^{\pi} q_\mu(z)\overline{q_\nu(z)}\,d\phi, \qquad z = e^{i\phi}. \tag{20}$$

According as n is even or odd, Legendre polynomials are even or odd functions, so that $K_{\mu\nu}=0$ if $\mu-\nu$ is odd. Furthermore, the term z^{n-2k} in $P_n(z)$ has the sign $(-1)^k$; consequently $K_{\mu\nu}$ is real and has the sign $(-1)^{(\mu-\nu)/2} = e^{i\pi(\mu-\nu)/2}$ if $\mu-\nu$ is even. Thus the maximum problem (19) can be separated into the two independent problems

$$\begin{cases} \max \sum^{(1)}_{\mu,\nu=0}{}^{n} K_{\mu\nu} x_\mu \bar{x}_\nu, & \sum_{\nu\equiv n} |x_\nu|^2 \leqq 1, \\ \max \sum^{(2)}_{\mu,\nu=0}{}^{n} K_{\mu\nu} x_\mu \bar{x}_\nu, & \sum_{\nu\equiv n-1} |x_\nu|^2 \leqq 1, \end{cases} \tag{21}$$

where, in $\sum^{(1)}$, $\mu\equiv\nu\equiv n \pmod 2$, while, in $\sum^{(2)}$, $\mu\equiv\nu\equiv n-1 \pmod 2$, and $\pi\lambda_n^{-1}$ is the *greater* of the corresponding maximum values.

2. We now propose to study some asymptotic properties of the integrals (20) using the asymptotic formula of Heine [2, p. 174]

$$P_n(z) \cong \frac{(z+(z^2-1)^{1/2})^{n+1/2}}{(2\pi n)^{1/2}(z^2-1)^{1/4}}. \tag{22}$$

Here z lies in the complex plane cut along the segment $(-1, +1)$. The ratio of the expressions in (22) tends to 1 uniformly in every closed region having no points in common with the segment $(-1, +1)$.

Let ω be a fixed positive integer, $n \to \infty$. We confine ourselves to the values of μ and ν characterized by the conditions

$$n - \omega \leqq \mu, \nu \leqq n, \qquad \mu - \nu \text{ even}. \tag{23}$$

Denoting again by γ, as in Part I, the arc $\pi/2-\epsilon \leqq \phi \leqq \pi/2+\epsilon$ and using (16), we have for $z = e^{i\phi} = ie^{i\phi'}$,

$$\begin{aligned}
&\int_\gamma q_\mu(z)\overline{q_\nu(z)}\,d\phi \\
&\quad\cong (\mu+\tfrac{1}{2})^{1/2}(\nu+\tfrac{1}{2})^{1/2}\int_\gamma \frac{(z+(z^2-1)^{1/2})^{\mu+1/2}\overline{(z+(z^2-1)^{1/2})}^{\nu+1/2}}{(2\mu\pi)^{1/2}(2\nu\pi)^{1/2}\,|\,(z^2-1)^{1/4}\,|^2}\,d\phi \\
&\quad\cong \exp\frac{i\pi(\mu-\nu)}{2}\;\frac{2^{-1/2}(2^{1/2}+1)^{\mu+\nu+1}}{2\pi} \\
&\qquad\cdot\int_{-\epsilon}^{\epsilon}\exp\left\{\left(2^{-1/2}i\phi' - 2^{-1/2}\frac{\phi'^2}{4}+\cdots\right)(\mu+\tfrac{1}{2})\right. \\
&\qquad\qquad\left.+\left(-2^{-1/2}i\phi' - 2^{-1/2}\frac{\phi'^2}{4}+\cdots\right)(\nu+\tfrac{1}{2})\right\}d\phi' \\
&\quad\cong (-1)^{(\mu-\nu)/2}\;\frac{2^{-1/2}(2^{1/2}+1)^{\mu+\nu+1}}{2\pi} \\
&\qquad\cdot\int_{-\epsilon}^{\epsilon}\exp\left\{2^{-1/2}i(\mu-\nu)\phi' - 2^{-1/2}(\mu+\nu+1)\frac{\phi'^2}{4}+\cdots\right\}d\phi' \\
&\quad\cong (-1)^{(\mu-\nu)/2}\;\frac{2^{-1/2}(2^{1/2}+1)^{\mu+\nu+1}}{2\pi}\int_{-\epsilon}^{\epsilon}\exp\left\{-2^{-1/2}(\mu+\nu+1)\frac{\phi'^2}{4}\right\}d\phi' \\
&\quad\cong (-1)^{(\mu-\nu)/2}\;\frac{2^{-1/2}(2^{1/2}+1)^{\mu+\nu+1}}{2\pi}\left\{2^{-1/2}\,\frac{\mu+\nu+1}{4}\right\}^{-1/2}\pi^{1/2} \\
&\quad\cong (-1)^{(\mu-\nu)/2}(2\pi 2^{1/2})^{-1/2}n^{-1/2}(2^{1/2}+1)^{\mu+\nu+1},
\end{aligned}$$

and therefore

$$K_{\mu\nu} \cong (-1)^{(\mu-\nu)/2}2^{1/4}\pi^{-1/2}n^{-1/2}(2^{1/2}+1)^{\mu+\nu+1}. \tag{24}$$

Using Schwarz's inequality we obtain *for all values of* μ *and* ν

$$|K_{\mu\nu}| \leqq K_{\mu\mu}^{1/2} K_{\nu\nu}^{1/2}. \tag{25}$$

On the other hand, for the values of μ and ν satisfying (23) on account of (24),

$$K_{\mu\nu} \cong (-1)^{(\mu-\nu)/2} K_{\mu\mu}^{1/2} K_{\nu\nu}^{1/2}. \tag{25'}$$

3. Now if $\{\epsilon_\nu\}$ denotes a suitable sequence with positive terms $\to 0$,

$$\begin{aligned}
\sum_{\mu,\nu=0}^{n}{}^{(1)} K_{\mu\nu} x_\mu \bar{x}_\nu &\leqq \sum |K_{\mu\nu}|\,|x_\mu|\,|x_\nu| \leqq \sum K_{\mu\mu}^{1/2} K_{\nu\nu}^{1/2} |x_\mu|\,|x_\nu| \\
&= \left\{ \sum_{\nu \equiv n} K_{\nu\nu}^{1/2} |x_\nu| \right\}^2 \leqq \sum_{\nu \equiv n} K_{\nu\nu} \\
&\leqq 2^{1/4}\pi^{-1/2}(2^{1/2}+1)^{2n+1} \sum_{\nu \equiv n} (1+\epsilon_\nu)\nu^{-1/2}(2^{1/2}+1)^{2\nu-2n} \qquad (26) \\
&\cong 2^{1/4}\pi^{-1/2}(2^{1/2}+1)^{2n+1} n^{-1/2} \sum_{k=0}^{\infty} (2^{1/2}+1)^{-4k} \\
&= 2^{-9/4}\pi^{-1/2} n^{-1/2} (2^{1/2}+1)^{2n+3}.
\end{aligned}$$

On the other hand we put

$$\begin{aligned}
&x_\nu = \sigma e^{i\pi\nu/2} K_{\nu\nu}^{1/2} \quad \text{if } \nu \equiv n, \qquad n-\omega \leqq \nu \leqq n, \\
&x_\nu = 0 \qquad\qquad \text{if } \nu \equiv n, \qquad\qquad \nu < n-\omega.
\end{aligned} \tag{27}$$

Here the positive number σ has to be determined according to the condition $\sum_{\nu \equiv n} |x_\nu|^2 = \sigma^2 \sum_{n-\omega \leqq \nu \leqq n;\, \nu \equiv n} K_{\nu\nu} = 1$. The form in question becomes then

$$\sum_{\substack{\mu,\nu=n-\omega \\ \mu \equiv \nu \equiv n}}^{n} K_{\mu\nu}\sigma^2 \exp\left[i\pi(\mu-\nu)/2\right] K_{\mu\mu}^{1/2} K_{\nu\nu}^{1/2}.$$

In view of (25′) we obtain for this expression

$$\begin{aligned}
&\cong \sigma^2 \sum K_{\mu\mu} K_{\nu\nu} = \sigma^2 \Big\{ \sum K_{\nu\nu} \Big\}^2 = \sum_{\substack{\nu=n-\omega \\ \nu \equiv n}}^{n} K_{\nu\nu} \\
&\cong 2^{1/4}\pi^{-1/2} n^{-1/2} (2^{1/2}+1)^{2n+1} \sum_{\substack{\nu=n-\omega \\ \nu \equiv n}}^{n} (2^{1/2}+1)^{2\nu-2n}.
\end{aligned}$$

Since the number ω is arbitrarily large the asymptotic value of the first maximum (21) is given by the right-hand member in (26). The second maximum is obtained by replacing here n by $n-1$; the result is of lower order than the preceding one which therefore furnishes the asymptotic evaluation of (19).

PART III

The treatment of the form (7) does not present new difficulties. Using notations analogous to those employed in Part II, it is sufficient to point out only the essential differences.

The orthogonal polynomials belonging to the interval (0, 1) are

$$q_n(t) = (2n+1)^{1/2}P_n(2t-1) = (2n+1)^{1/2}\sum_{k=0}^{n}\frac{P_n^{(k)}(-1)}{k!}(2t)^k. \tag{28}$$

Hence we obtain

$$K_{\mu\nu} = (2\mu+1)^{1/2}(2\nu+1)^{1/2}\int_{-\pi}^{\pi}P_\mu(2z-1)\overline{P_\nu(2z-1)}d\phi, \qquad z = e^{i\phi}. \tag{29}$$

Since $\operatorname{sgn} P_n^{(k)}(-1) = (-1)^{n-k}$ we easily see that $K_{\mu\nu}$ is real and has the sign $(-1)^{\mu-\nu}$. The inequality (25) holds again.

In this case, $z = e^{i\phi}$,

$$\max\left|2z-1+[(2z-1)^2-1]^{1/2}\right| = 3+2(2^{1/2})$$

and this value is attained for $z=-1$. Thus the main part of our integrals is given by the arc $\pi-\epsilon \leqq \phi \leqq \pi+\epsilon$. We need therefore, the development of

$$\log\left\{2z-1+[(2z-1)^2-1]^{1/2}\right\} = 2\log\left\{z^{1/2}+((z^{1/2})^2-1)^{1/2}\right\} \tag{30}$$

if $z = e^{i\phi} = -e^{i\phi'}$ and ϕ' is small. Since $z^{1/2} = ie^{i\phi'/2}$ this can be readily calculated from (16) in the form

$$i\pi + 2\log(2^{1/2}+1) + 2^{-1/2}i\phi' - 2^{1/2}\frac{\phi'^2}{16} + \cdots. \tag{30'}$$

This yields under the condition

$$n-\omega \leqq \mu,\ \nu \leqq n \tag{31}$$

the asymptotic formula

$$\begin{aligned} K_{\mu\nu} &\cong (2\mu+1)^{1/2}(2\nu+1)^{1/2}\int_{\pi-\epsilon}^{\pi+\epsilon}P_\mu(2z-1)\overline{P_\nu(2z-1)}d\phi \\ &\cong (2\mu+1)^{1/2}(2\nu+1)^{1/2} \\ &\cdot\int_{\pi-\epsilon}^{\pi+\epsilon}\frac{\left\{2z-1+((2z-1)^2-1)^{1/2}\right\}^{\mu+1/2}\left\{\overline{2z-1+((2z-1)^2-1)^{1/2}}\right\}^{\nu+1/2}}{(2\pi\mu)^{1/2}(2\pi\nu)^{1/2}\left|((2z-1)^2-1)^{1/4}\right|^2}d\phi \\ &\cong (-1)^{\mu-\nu}(2\pi 2^{1/2})^{-1}(2^{1/2}+1)^{2\mu+2\nu+2} \\ &\cdot\int_{-\epsilon}^{+\epsilon}\exp\left\{2^{-1/2}i(\mu-\nu)\phi' - 2^{1/2}(\mu+\nu+1)\frac{\phi'^2}{16}\right\}d\phi' \end{aligned} \tag{32}$$

$$\cong (-1)^{\mu-\nu}(2\pi 2^{1/2})^{-1}(2^{1/2}+1)^{2\mu+2\nu+2}\left\{\frac{2^{1/2}(\mu+\nu+1)}{16}\right\}^{-1/2}\pi^{1/2}$$

$$\cong (-1)^{\mu-\nu}2^{-1/4}\pi^{-1/2}n^{-1/2}(2^{1/2}+1)^{2\mu+2\nu+2}.$$

For the values (31) the asymptotic formula (25′) holds again, with $(-1)^{(\mu-\nu)/2}$ replaced by $(-1)^{\mu-\nu}$.

Consequently, by means of an argument similar to that used in Part II, we obtain for

$$2\pi\lambda_n^{-1} = \max \sum K_{\mu\nu}x_\mu\bar{x}_\nu, \qquad \sum |x_\nu|^2 \leqq 1,$$

$$2\pi\lambda_n^{-1} \cong \sum_{\nu=0}^{n} K_{\nu\nu} \cong 2^{-1/4}\pi^{-1/2}n^{-1/2}(2^{1/2}+1)^{4n+2}\sum_{k=0}^{\infty}(2^{1/2}+1)^{-4k}$$

$$= 2^{-11/4}\pi^{-1/2}n^{-1/2}(2^{1/2}+1)^{4n+4}.$$

Part IV

1. Introducing the normalized Hermite polynomials $q_n(t)$ belonging to the weight function e^{-t^2}, we see that the smallest characteristic value λ_n of (9) is given by

$$2\pi\lambda_n^{-1} = \max \sum K_{\mu\nu}x_\mu\bar{x}_\nu, \qquad \sum |x_\nu|^2 \leqq 1,$$

where

$$K_{\mu\nu} = \int_{-\pi}^{\pi} q_\mu(z)\overline{q_\nu(z)}d\phi, \qquad z = e^{i\phi}. \tag{33}$$

Hermite polynomials can be represented easily in terms of Laguerre polynomials with the parameter values $\pm\frac{1}{2}$. Using the notation from the book by Pólya-Szegö [**5,** vol. 2, p. 94, problem 99] we obtain [cf. loc. cit., p. 95, problem 102]

$$q_n(t) = \begin{cases}(\Gamma(m+\frac{1}{2}))^{-1/2}(\Gamma(m+1))^{1/2}L_m^{(-1/2)}(t^2) & \text{if } n = 2m,\\ (\Gamma(m+\frac{3}{2}))^{-1/2}(\Gamma(m+1))^{1/2}tL_m^{(1/2)}(t^2) & \text{if } n = 2m+1.\end{cases} \tag{34}$$

Consequently $K_{\mu\nu}=0$ if $\mu-\nu$ is odd. For $\mu-\nu$ even, $K_{\mu\nu}$ is real and positive.

According to the asymptotic formula of Laguerre polynomials due to Perron [cf. **4,** p. 77, (48)]

$$q_n(z) \cong \eta(2\pi^{1/2})^{-1}\left(\frac{n}{2}\right)^{-1/4}\exp\left[z^2/2 - iz(2n)^{1/2}\right], \tag{35}$$

where $\eta=1$ if n is even, $\eta=i$ if n is odd. Here z is supposed to have a positive imaginary part. The essential part of $K_{\mu\nu}$ is given again by an arc γ of the

type of Part II.* Let ω be a fixed positive number; for large values of n and for

$$n - \omega n^{1/2} \leqq \mu, \nu \leqq n, \qquad \mu - \nu \text{ even}, \tag{36}$$

we have

$$\begin{aligned}\int_\gamma q_\mu(z)\overline{q_\nu(z)}d\phi \\ \cong (4\pi e)^{-1}\left(\frac{\mu}{2}\right)^{-1/4}\left(\frac{\nu}{2}\right)^{-1/4}\int_\gamma \exp\left[-iz(2\mu)^{1/2} + i\bar{z}(2\nu)^{1/2}\right]d\phi.\end{aligned} \tag{37}$$

Putting $z = ie^{i\phi'}$ we obtain

$$\begin{aligned}&(4\pi e)^{-1}\left(\frac{n}{2}\right)^{-1/2}\int_{-\epsilon}^{\epsilon}\exp\left\{\left(1 + i\phi' - \frac{\phi'^2}{2}\right)(2\mu)^{1/2}\right. \\ &\qquad\left. + \left(1 - i\phi' - \frac{\phi'^2}{2}\right)(2\nu)^{1/2}\right\}d\phi' \\ &\cong (4\pi e)^{-1}\left(\frac{n}{2}\right)^{-1/2}\exp\left[(2\mu)^{1/2} + (2\nu)^{1/2}\right] \\ &\qquad\cdot\int_{-\epsilon}^{\epsilon}\exp\left[-\tfrac{1}{2}((2\mu)^{1/2} + (2\nu)^{1/2})\phi'^2\right]d\phi' \qquad (37') \\ &\cong (4\pi e)^{-1}\left(\frac{n}{2}\right)^{-1/2}\exp\left[(2\mu)^{1/2} + (2\nu)^{1/2}\right]\left\{\tfrac{1}{2}((2\mu)^{1/2} + (2\nu)^{1/2})\right\}^{-1/2}\pi^{1/2} \\ &\cong 2^{-7/4}\pi^{-1/2}e^{-1}n^{-3/4}\exp\left[(2\mu)^{1/2} + (2\nu)^{1/2}\right]\end{aligned}$$

because $\mu^{1/2} - \nu^{1/2}$ remains bounded. Hence, for the values of μ, ν defined by (36)

$$K_{\mu\nu} \cong 2^{-3/4}\pi^{-1/2}e^{-1}n^{-3/4}\exp\left[(2\mu)^{1/2} + (2\nu)^{1/2}\right]. \tag{38}$$

Hence we get as in Part II

$$\begin{aligned}2\pi\lambda_n^{-1} &\cong \sum_{\nu=n_1} K_{\nu\nu} \cong 2^{-3/4}\pi^{-1/2}e^{-1}n^{-3/4}\sum_{\nu=n_1}\exp\left[2(2\nu)^{1/2}\right] \\ &\cong 2^{-3/4}\pi^{-1/2}e^{-1}n^{-3/4}\cdot\frac{1}{2}\int_1^n \exp\left[2(2x)^{1/2}\right]dx \\ &= 2^{-3/4}\pi^{-1/2}e^{-1}n^{-3/4}\int_1^{n^{1/2}} x\exp\left[2(2^{1/2})x\right]dx \cong 2^{-9/4}\pi^{-1/2}e^{-1}n^{-1/4}\exp\left[2(2n)^{1/2}\right].\end{aligned}$$

Thus the first statement (11) is established.

2. Dealing with the form (10) we use Laguerre polynomials $q_n(t) = L_n(t)$.

* Near the real axis, i.e., in the neighborhood of the points $z = \pm 1$, the formula (35) cannot be used. Here the fact has to be regarded that $|q_n(x+iy)|$, x and y real, is monotonically *increasing* if $|y|$ increases. This is a consequence of the reality of the zeros of $q_n(z)$.

Here $K_{\mu\nu}$ is always real and positive. Excluding the case in which z is real and non-negative, we have [Perron, **4**, loc. cit.]

$$q_n(z) \cong (2\pi^{1/2})^{-1}(-z)^{-1/4}e^{z^{1/2}n^{-1/4}} \exp[2(-zn)^{1/2}]$$

The essential part of $K_{\mu\nu}$ is given now by the contribution of a neighborhood of $z=-1$ (cf. the footnote on p. 460).

Restricting μ and ν as in (36) but without the condition regarding $\mu-\nu$ we obtain

$$K_{\mu\nu} \cong \int_{\pi-\epsilon}^{\pi+\epsilon} q_\mu(z)\overline{q_\nu(z)}d\phi$$

$$\cong (4\pi e)^{-1}\mu^{-1/4}\nu^{-1/4}\int_{\pi-\epsilon}^{\pi+\epsilon} \exp\left[2(-z\mu)^{1/2} + 2(-\bar{z}\nu)^{1/2}\right]d\phi.$$

Writing $z=-e^{i\phi'}$ we get

$$(4\pi e)^{-1}n^{-1/2}\int_{-\epsilon}^{+\epsilon} \exp\left\{2\mu^{1/2}\left(1+i\frac{\phi'}{2}-\frac{\phi'^2}{8}\right)+2\nu^{1/2}\left(1-i\frac{\phi'}{2}-\frac{\phi'^2}{8}\right)\right\}d\phi'$$

$$\cong (4\pi e)^{-1}n^{-1/2}\exp\left[2\mu^{1/2}+2\nu^{1/2}\right]\int_{-\epsilon}^{\epsilon}\exp\left[-(\mu^{1/2}+\nu^{1/2})\phi'^2/4\right]d\phi'$$

$$\cong (4\pi e)^{-1}n^{-1/2}\exp\left[2\mu^{1/2}+2\nu^{1/2}\right]\left\{\frac{\mu^{1/2}+\nu^{1/2}}{4}\right\}^{-1/2}\pi^{1/2}$$

$$\cong 2^{-3/2}\pi^{-1/2}e^{-1}n^{-3/4}\exp\left[2\mu^{1/2}+2\nu^{1/2}\right],$$

therefore

$$2\pi\lambda_n^{-1} \cong \sum_{\nu=0}^{n} K_{\nu\nu} \cong 2^{-3/2}\pi^{-1/2}e^{-1}n^{-3/4}\int_1^n \exp\left[4x^{1/2}\right]dx \cong 2^{-1/2}\pi^{-1/2}e^{-1}n^{-1/4}\exp\left[4n^{1/2}\right].$$

Bibliography

1. M. Fekete, *Über den absoluten Betrag von Polynomen, welche auf einer Punktmenge gleichmässig beschränkt sind*, Mathematische Zeitschrift, vol. 26 (1927), pp. 324–344.

2. H. E. Heine, *Handbuch der Kugelfunktionen*, 2d edition, Berlin, 1878.

3. D. Hilbert, *Ein Beitrag zur Theorie des Legendre'schen Polynoms*, Acta Mathematica, vol. 18 (1894), pp. 155–159.

4. O. Perron, *Über das Verhalten einer ausgearteten hypergeometrischen Reihe bei unbegrenztem Wachstum eines Parameters*, Journal für Mathematik, vol. 151 (1921), pp. 63–78.

5. G. Pólya-G. Szegö, *Aufgaben und Lehrsätze aus der Analysis*, 2 volumes, Berlin, 1925.

6. I. Schur, *Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen*, Journal für Mathematik, vol. 140 (1911), pp. 1–28.

7. G. Szegö, *A Hankel-féle formákról*, Matematikai és Természettudományi Értesítö, vol. 36 (1918), pp. 497–538.

8. G. Szegö, *Über orthogonale Polynome, die zu einer gegebenen Kurve der komplexen Ebene gehören*, Mathematische Zeitschrift, vol. 9 (1921), pp. 218–270.

Washington University,
St. Louis, Mo.

ON A CLASSIFICATION OF INTEGRAL FUNCTIONS BY MEANS OF CERTAIN INVARIANT POINT PROPERTIES*

BY

R. D. CARMICHAEL, W. T. MARTIN,† AND M. T. BIRD

1. **Introduction.** In the investigation of linear differential equations of infinite order and of certain other types of linear functional equations, the authors of this paper have severally had occasion to employ certain point properties of integral functions which remain unaltered in passing from one regular point to another. These invariant point properties are of considerable interest in themselves, and the theorems to which they give rise extend classic results in the theory of integral functions. In this paper we lay the foundations of a general theory of these invariant point properties of integral functions and present the principal results concerning them which we have severally drawn upon in our recent investigations (not yet published).

2. **On I-sequences.** Throughout the paper the symbol $\{t_n\}$ will denote an infinite sequence of positive numbers $t_0, t_1, t_2, \cdots$. We define associated positive-termed sequences $\{\lambda_n\}$ and $\{\rho_n\}$ by means of the relations

$$t_0 = \lambda_0 = \rho_0, \qquad t_n = \lambda_1\lambda_2 \cdots \lambda_n = (\rho_n)^n, \qquad (n = 1, 2, 3, \cdots).$$

An infinite sequence $\{t_n\}$ will be called an I-sequence when and only when

(a) it consists of positive numbers having the property that for every single-valued analytic function $F(x)$ we have

$$\limsup_{n=\infty} \left| t_n F^{(n)}(x_0)/n! \right|^{1/n} = \limsup_{n=\infty} \left| t_n F^{(n)}(x_1)/n! \right|^{1/n} \tag{1}$$

for every two finite regular points x_0 and x_1 of $F(x)$;

(b) there exists at least one analytic function $F(x)$ for which the superior limits in (1) have a value different from zero.

The property of $F(x)$ expressed by (1) is an invariant point property of of the function.

If $\{t_n\}$ is an I-sequence and if $F(x)$ is a particular analytic function for which the superior limits in (1) have the value q $(0 \leqq q \leqq \infty)$ then we shall say that $F(x)$ is of type $(t_0, t_1, t_2, \cdots; q)$, or, more briefly, of type $(t_n; q)$. When q is finite we shall speak of the type as normal.

* Presented to the Society, April 11, 1936; received by the editors February 21, 1936.

† National Research Fellow.

It is obvious that every polynomial is of type $(t_n; 0)$ for every I-sequence $\{t_n\}$. Hence any particular I-sequence serves to separate all analytic functions into (two or more) types or classes.

If $\{t_n\}$ is an I-sequence it is clear that $\{\tau_n\}$ is also an I-sequence provided that its elements are positive and that the limit

$$\lim_{n=\infty} (\tau_n/t_n)^{1/n} \tag{2}$$

exists as a finite quantity p different from zero. Then if $F(x)$ is of type $(t_n; q)$ it is also obviously of type $\{\tau_n; pq\}$.

Let us put

$$\limsup_{n=\infty} (t_n)^{1/n} = \rho, \qquad \liminf_{n=\infty} (t_n)^{1/n} = \sigma, \qquad (0 \leqq \sigma \leqq \rho \leqq \infty).$$

We shall first prove that ρ is necessarily infinite for every I-sequence $\{t_n\}$. For, if we assume ρ to be finite, then for the particular function $F(x) = 1/(1-x)$ relation (1) leads to the equation

$$\rho \left| 1 - x_0 \right|^{-1} = \rho \left| 1 - x_1 \right|^{-1}$$

for all finite x_0 and x_1 different from unity; and this is impossible unless $\rho = 0$. But if $\rho = 0$ it is easy to see that condition (b) in the definition of an I-sequence is not satisfied.

When $\sigma \neq 0$ we denote by $E(x)$ the function

$$E(x) = \sum_{\nu=0}^{\infty} \frac{x^\nu}{t_\nu}, \tag{3}$$

the indicated power series converging for $|x| < \sigma$. Taking $x_0 = 0$ and x_1 such that $|x_1| < \sigma$ we have from (1) the relation

$$1 = \limsup_{n=\infty} \left| t_n E^{(n)}(x_1)/n! \right|^{1/n}$$

provided that $\{t_n\}$ is an I-sequence, as we now assume. Hence $E(x)$ is analytic in the interior of a circle about x_1 as center and of radius σ. But this is not possible for every x_1 satisfying the indicated condition unless σ is infinite, since if σ is finite $E(x)$ has a singularity on the circle $|x| = \sigma$.

On combining the results obtained, we have the following theorem:

THEOREM I. *In order that $\{t_n\}$ shall be an I-sequence, it is necessary that one or the other of the two following conditions shall be satisfied*:

$$\lim_{n=\infty} (t_n)^{1/n} = \infty; \tag{4}$$

$$\limsup_{n=\infty} (t_n)^{1/n} = \infty \quad \textit{and} \quad \liminf_{n=\infty} (t_n)^{1/n} = 0. \tag{5}$$

An I-sequence $\{t_n\}$ for which condition (4) [(5)] is satisfied will be called an A-sequence [an α-sequence]. We shall be concerned with A-sequences. We have not been able to determine whether α-sequences exist; we know some of their properties if they do exist; but we shall not present those properties in this paper.

3. **On A-sequences.** For every A-sequence $\{t_n\}$ the function $E(x)$ defined in (3) is an integral function. It is of normal type $(t_n; 1)$. If $\{t_n\}$ is an A-sequence and a given analytic function $F(x)$ is of *normal* type $\{t_n; q\}$, then $F(x)$ is clearly an integral function. Hence the usefulness of A-sequences will be confined mainly, if not entirely, to the class of integral functions. It is obvious that (4) is necessary as well as sufficient to meet the requirement that *every* function of *normal* type $(t_n; q)$ shall be an integral function.

We shall now develop necessary and sufficient conditions that $\{t_n\}$ shall be an A-sequence. We suppose condition (4) to be satisfied. Then $E(x)$ is an integral function. In order that $\{t_n\}$ shall be an A-sequence, it is necessary that for every positive number a we have

$$\limsup_{n=\infty} (t_n E^{(n)}(a)/n!)^{1/n} = 1.$$

This is obtained from (1) by taking $x_0 = a$, $x_1 = 0$, and $F(x) \equiv E(x)$. But it is easily seen that the inferior limit of this limitand is not less than unity since for each value of n the limitand itself is greater than unity. Consequently a necessary condition that a sequence $\{t_n\}$ of positive terms satisfying (4) shall be an A-sequence is that we shall have

$$\lim_{n=\infty} (t_n E^{(n)}(a)/n!)^{1/n} = 1 \tag{6}$$

for every positive value of a.

Let us now suppose that $\{t_n\}$ is a sequence of positive terms satisfying conditions (4) and (6); and let $F(x)$ be a single-valued analytic function which is regular at the finite point x_0 for which

$$\limsup_{n=\infty} | t_n F^{(n)}(x_0)/n! |^{1/n}$$

has the finite value $q = q(x_0)$. Then $F(x)$ is an integral function. Moreover, for every positive number ϵ there exists a positive number M_ϵ such that

$$| F^{(\mu)}(x_0) | < M_\epsilon (q + \epsilon)^\mu \mu!/t_\mu, \qquad (\mu = 0, 1, 2, \cdots).$$

Then from the power series expansion of $F^{(n)}(x)$ in powers of $x - x_0$ we have

$$\begin{aligned} | F^{(n)}(x) | &< M_\epsilon (q + \epsilon)^n \sum_{\nu=0}^{\infty} \frac{(n+\nu)!}{\nu!} \frac{1}{t_{n+\nu}} (q + \epsilon)^\nu | x - x_0 |^\nu \\ &= M_\epsilon (q + \epsilon)^n E^{(n)}[(q + \epsilon) | x - x_0 |]. \end{aligned} \tag{7}$$

Then from (6) with $a=(q+\epsilon)|x-x_0|$ and from the arbitrariness of the ϵ we see that

$$\limsup_{n=\infty} |t_n F^{(n)}(x)/n!|^{1/n}$$

is not greater than q. If this superior limit were less than q for some x_1, we could start from this x_1 and prove similarly that it is less than q for $x=x_0$, contrary to hypothesis. Hence its value is q.

This enables us to conclude to the following theorem:

THEOREM II. *A necessary and sufficient condition that a sequence $\{t_n\}$ of positive terms $t_0, t_1, t_2, \cdots$ shall be an A-sequence is that it shall satisfy condition* (4) *and the condition*

$$\lim_{n=\infty}\left(\sum_{\nu=0}^{\infty}\frac{t_n}{t_{\nu+n}}\frac{(\nu+n)!}{\nu!n!}a^\nu\right)^{1/n}=1 \tag{8}$$

for every positive value of a.

Equation (8) is merely another form of (6). An equivalent condition is set forth in the following corollary, which we now prove:

COROLLARY 1. *Condition* (8) *in the theorem may be replaced by the following equivalent condition*:

$$\lim_{n=\infty}\left(\sum_{\nu=0}^{\infty}\frac{t_n}{t_{\nu+n}}a^\nu\right)^{1/n}=1. \tag{9}$$

Since the binomial theorem implies that

$$(1+\delta)^{n+\nu}\geqq\frac{(\nu+n)!}{\nu!n!}\delta^\nu$$

for every positive δ and for non-negative integers n and ν, it follows that

$$\frac{(\nu+n)!}{\nu!n!}\leqq(1+\delta)^n\left(1+\frac{1}{\delta}\right)^\nu \tag{10}$$

for every positive δ. Hence we have

$$\sum_{\nu=0}^{\infty}\frac{t_n}{t_{\nu+n}}\frac{(\nu+n)!}{\nu!n!}a^\nu\leqq(1+\delta)^n\sum_{\nu=0}^{\infty}\frac{t_n}{t_{\nu+n}}\left(1+\frac{1}{\delta}\right)^\nu a^\nu.$$

From this relation and from (9) with a replaced by $a(1+1/\delta)$ it follows that (9) implies (8). That (8) implies (9) follows readily from the fact that the limitand in (9) lies between 1 and the limitand in (8). Hence the corollary is established.

We shall now prove the following two corollaries of Theorem II:

COROLLARY 2. *Any one of the following is a necessary condition that the sequence* $\{t_n\}$ *be an A-sequence*:

(a) $\lim_{n=\infty} \rho_n = \infty$, (b) $\limsup_{n=\infty} \lambda_n = \infty$,

(c) $\liminf_{n=\infty} (\lambda_n)^{1/n} \geqq 1$, (d) $\lim_{n=\infty} (\lambda_{n+1}\lambda_{n+2} \cdots \lambda_{2n})^{1/n} = \infty$.

COROLLARY 3. *Either one of the following is a sufficient condition that the positive-termed sequence* $\{t_n\}$ *be an A-sequence*:

(a) $\lim_{n=\infty} \lambda_n = \infty$, (b) $\rho_{n+1} \geqq \rho_n$, $\lim_{n=\infty} \rho_n = \infty$.

In Corollary 2 (a) we have merely a repetition of Condition (4), and 2 (b) is an immediate consequence of 2 (a). Corollaries 2 (c) and 2 (d) follow immediately from a consideration of the second and the $(n+1)$th terms in the series in (9). When the condition in Corollary 3 (a) is satisfied we have equation (4). When the same condition is satisfied, then for any K whatever greater than a we have

$$1 < 1 + \sum_{\nu=1}^{\infty} \frac{(n+\nu)!}{n!\nu!} \frac{a^\nu}{\lambda_{n+1} \cdots \lambda_{n+\nu}} < \sum_{\nu=0}^{\infty} \frac{(n+\nu)!}{n!\nu!} \frac{a^\nu}{K^\nu} = \left(1 - \frac{a}{K}\right)^{-n-1}$$

for all n greater than a given value depending on K. Hence Condition (8) is satisfied. Thence from Theorem II we conclude to the truth of Corollary 3(a). This corollary in turn implies Corollary 3 (b).

By aid of Corollary 3 it is easy to show that the sequence $\{t_n\}$ is an A-sequence for each of the following definitions of t_n:

(11) $t_n = (n!)^t$ or $t_n = e^{nt+1}$ or $t_n = n^{tn}$ (except that $t_0 = 1$), $\quad 0 < t < \infty$;

(12) $t_n = e^{n(\log n)^t}$, $0 < t < \infty$, $n > 0$, $t_0 = 1$;

(13) $t_n = (\log n)^{tn}$ or $t_n = n^{n[1-(\log n)^{-t}]}$, $0 < t < \infty$, $n \geqq 2$, $t_0 = t_1 = 1$;

(14) $t_n = e^{tnl_k(n)}$ when $l_k(n)$ is real but otherwise $t_n = 1$, where, in the last case, we have

$$0 < t < \infty, \; l_1(n) = \log n, \; l_\mu(n) = \log [l_{\mu-1}(n)], \; \mu > 1.$$

Functions of normal type $(n!; q)$ are the functions of exponential type.* Functions of type $((n!)^t; q)$, with $0 < t < \infty$, and $0 < q < \infty$, are the functions of order $1/t$ and normal type $tq^{1/t}$ in accordance with the classic terminology† of the theory of integral functions.

* See, for instance, R. D. Carmichael, Bulletin of the American Mathematical Society, vol. 40 (1934), pp. 241-261.

† See, for instance, A. Pringsheim, Mathematische Annalen, vol. 58 (1904), pp. 257-342.

We shall now prove the following theorem:

THEOREM III. *Let* $\{t_n\}$ *be an A-sequence. Let* $\{t_n'\}$, *with* $t_0' = \lambda_0'$, $t_n' = \lambda_1'\lambda_2' \cdots \lambda_n'$ *for* $n = 1, 2, 3, \cdots$, *be any infinite sequence of positive numbers such that a positive constant M exists so that* $\lambda_n' \geqq M\lambda_n$ *for* $n = 1, 2, \cdots$. *Then* $\{t_n'\}$ *is an A-sequence.*

Since t_n satisfies Condition (4) it is evident that t_n' satisfies a like condition. For positive n and positive a we have

$$1 < \sum_{\nu=0}^{\infty} \frac{(n+\nu)!}{n!\nu!} \frac{a^\nu}{\lambda_{n+1}' \cdots \lambda_{n+\nu}'} \leqq \sum_{\nu=0}^{\infty} \frac{(n+\nu)!}{n!\nu!} \frac{(a/M)^\nu}{\lambda_{n+1} \cdots \lambda_{n+\nu}}.$$

Hence Condition (8) on t_n implies the corresponding condition on t_n'. Thence it follows that Theorem III is implied by Theorem II.

4. **Classification of integral functions.** With respect to the general question of the existence of A-sequences suitable for classifying transcendental integral functions, we prove the following theorem:

THEOREM IV. *For every transcendental integral function* $F(x)$ *there exists an associated A-sequence* $\{t_n\}$ *such that* $F(x)$ *is of type* $(t_n; 1)$ *and the corresponding* ρ_n *is monotonic non-decreasing.*

We define the required sequence $\{t_n\}$ as follows. For positive n, let t_n be a positive number such that $t_n^{-1/n}$, or $1/\rho_n$, is the greatest of the numbers

$$\left| F^{(\mu)}(0)/\mu! \right|^{1/\mu}, \qquad (\mu = n, n+1, \cdots),$$

such greatest numbers* obviously existing since $F(x)$ is a transcendental *integral* function. Put $\rho_0 = \rho_1$. From Corollary 3(b) to Theorem II it follows readily that such a sequence $\{t_n\}$ is an A-sequence, since $t_n^{-1/n}$ is clearly a monotonic non-increasing function of n approaching zero as n becomes infinite, whence the corresponding ρ_n is monotonic non-decreasing and becomes infinite with n. Thence it is easy to see that $F(x)$ is of type $(t_n; 1)$ and that the theorem follows as stated.

The totality of A-sequences $\{t_n\}$ associated with the totality of transcendental integral functions in the manner indicated in the foregoing proof contains just those A-sequences for which ρ_n is monotonic non-decreasing, as one sees from a consideration of the particular function

$$F(x) = 1 + \rho_1^{-1}x + \rho_2^{-2}x^2 + \cdots.$$

5. **Bounds on functions of normal type** $(t_n; q)$ **for** A**-sequences** $\{t_n\}$. Equa-

* Such sequences $\{1/\rho_n\}$ have been considered by G. Faber, Mathematische Annalen, vol. 70 (1911), pp. 48–68; see p. 51.

tion (7) affords an upper bound to $|F^{(n)}(x)|$ in terms of the nth derivative of $E(x)$ when $F(x)$ is a function of normal type $(t_n; q)$ and $\{t_n\}$ is an A-sequence. It is sometimes useful to have such an upper bound in terms of $E(x)$ itself rather than in terms of its derivatives. In this connection we prove the following theorem:

THEOREM V. *Let $\{t_n\}$ be an A-sequence such that the corresponding ρ_n is monotonic non-decreasing. Then if $F(x)$ is of normal type $(t_n; q)$, we have for every positive δ and every positive ϵ an $M_\epsilon(x_0)$ depending on ϵ and x_0 alone such that*

$$(15) \quad |t_n F^{(n)}(x)/n!| < M_\epsilon(x_0)(q+\epsilon)^n(1+1/\delta)^n E[(q+\epsilon)(1+\delta)|x-x_0|]$$

for every finite x and x_0.

From (7) and (10) we have

$$|t_n F^{(n)}(x)/n!| < M_\epsilon(x_0)(q+\epsilon)^n(1+1/\delta)^n \sum_{\nu=0}^{\infty} \frac{t_n}{t_{n+\nu}} (q+\epsilon)^\nu(1+\delta)^\nu |x-x_0|^\nu.$$

But since ρ_n is now monotonic we have for $n\nu>0$ the relations

$$t_n t_\nu = \rho_n^n \rho_\nu^\nu \leqq \rho_{n+\nu}^n \rho_{n+\nu}^\nu = t_{n+\nu}, \text{ whence } t_n/t_{n+\nu} \leqq 1/t_\nu.$$

From these relations and the definition of $E(x)$ we have (15), and the theorem is proved.

6. **Type of the derived function.** We now prove the following theorem:

THEOREM VI. *If $F(x)$ is of type $(t_n; q)$, for a given A-sequence $\{t_n\}$, then $F'(x)$ is of type $(t_n; q_1)$ where $q_1 \leqq q$. When q is finite, a necessary and sufficient condition that $q_1 = q$ in all cases is that* $\lim_{n=\infty}(\lambda_n)^{1/n} = 1$.

We have

$$\limsup_{n=\infty} |t_n F^{(n+1)}(x_0)/n!|^{1/n} = \limsup_{n=\infty} |t_n F^{(n+1)}(x_0)/n!|^{1/(n+1)}$$

$$\leqq \limsup_{n=\infty} |t_{n+1} F^{(n+1)}(x_0)/(n+1)!|^{1/(n+1)} \cdot \limsup_{n=\infty} (1/\lambda_{n+1})^{1/(n+1)}.$$

Now apply Corollary 2(c) following Theorem II. The stated theorem follows readily.

7. **Representation of functions of normal type $(t_n; q)$.** We prove the following theorem:

THEOREM VII. *A necessary and sufficient condition that a function $F(x)$ be of normal type $(t_n; q)$, for a given A-sequence $\{t_n\}$, is that it be expressible in the form*

$$F(x) = \frac{1}{2\pi i}\int_C E(xz)a(z)dz, \tag{16}$$

where $a(z)$ is a function which is analytic in the ring $q<|z|\leqq f\leqq\infty$ and has a singularity on the circle $|z|=q$, and where C is a circle of radius $q+\epsilon$, $0<\epsilon<f-q$, about the point 0 as a center. Here $E(x)$ is the function defined in (3).

An easily verified representation of $F(x)$ in the form indicated is the following:

$$F(x) = \frac{1}{2\pi i}\int_C E(xz)\left(\sum_{n=0}^{\infty}\frac{t_nF^{(n)}(0)}{n!z^{n+1}}\right)dz.$$

Conversely, if $F(x)$ has the representation (16), then we may write

$$a(z) = \sum_{\nu=0}^{\infty}\frac{a_\nu}{z^{\nu+1}} + \sum_{\nu=0}^{\infty}b_\nu z^\nu, \qquad q<|z|\leqq f,$$

where $\limsup_{\nu=\infty}|a_\nu|^{1/\nu}=q$. Then from (16) we have

$$F(x) = \sum_{\nu=0}^{\infty}\frac{a_\nu x^\nu}{t_\nu},$$

whence it follows readily that $F(x)$ is of normal type $(t_n; q)$.

8. **On B-sequences.** Independently of the notion of A-sequence, we shall now define a class of sequences, which we call B-sequences, and we shall then prove that every B-sequence is an A-sequence.

Let $\{t_n\}$ be an infinite sequence of positive numbers $t_0, t_1, t_2, \cdots$ verifying equation (4). Define $E(x)$ as in (3). If $F(x)$ is a single-valued analytic function which is regular at $x=x_0$ and if we have

$$\limsup_{n=\infty}|t_nF^{(n)}(x_0)/n!|^{1/n} = q = q(x_0), \qquad (0\leqq q<\infty), \tag{17}$$

whence $F(x)$ is an integral function, then for every positive number ϵ there exists a constant $M_\epsilon(x_0)$, depending on ϵ and x_0, such that

$$|F(x)| < M_\epsilon(x_0)E[(q+\epsilon)|x-x_0|] \tag{18}$$

for every finite x, as we may see through a proof similar to that by which equation (7) was derived.

Let us suppose conversely that the sequence $\{t_n\}$ verifies (4) and is such that every integral function $F(x)$ having the property associated with (18) has also the property that

$$\limsup_{n=\infty}|t_nF^{(n)}(x_0)/n!|^{1/n} \leqq q. \tag{19}$$

Then, and only then, we shall say that $\{t_n\}$ is a B-sequence.

THEOREM VIII. *Every B-sequence is an A-sequence.*

Let $\{t_n\}$ be a B-sequence. Define $E(x)$ as in equation (3). We have

$$|E(x)| \leqq E(|x|) = \sum_{\nu=0}^{\infty} \frac{|x - x_0|^\nu}{t_\nu} \left|1 + \frac{x_0}{x - x_0}\right|^\nu$$
$$\leqq M_\epsilon(x_0) E[(1+\epsilon)|x - x_0|].$$

Then, since $\{t_n\}$ is a B-sequence, we have

$$\limsup_{n=\infty} |t_n E^{(n)}(x_0)/n!|^{1/n} \leqq 1.$$

Letting x_0 run over all positive values a we evidently have the condition associated with (8), since the limitand is never less than 1. Since t_n verifies (4), it follows from Theorem II that $\{t_n\}$ is an A-sequence, as was to be proved.

THEOREM IX. *A sufficient condition that a sequence $\{t_n\}$ of positive numbers $t_0, t_1, t_2, \cdots$ satisfying* (4) *shall be a B-sequence is that positive numbers r_n shall exist such that we have*

$$\lim_{n=\infty} \left\{ \sum_{\nu=0}^{\infty} \frac{r_n^\nu}{r_n^n} \frac{t_n}{t_\nu} \right\}^{1/n} = 1. \tag{20}$$

Let $F(x)$ be a single-valued analytic function verifying a relation of the form (18) where $\{t_n\}$ is a sequence verifying (4). Then $F(x)$ is an integral function, and we may write

$$\frac{F^{(n)}(x_0)}{n!} = \frac{1}{2\pi i} \int_{\Gamma_n} \frac{F(z)}{(z - x_0)^{n+1}} dz,$$

where Γ_n is a circle of radius $r_n/(q+\epsilon)$ about the point x_0 as a center. From (18) and the last relation we have

$$|t_n F^{(n)}(x_0)/n!| \leqq M_\epsilon(x_0)(q+\epsilon)^n t_n r_n^{-n} E(r_n) = M_\epsilon(x_0)(q+\epsilon)^n \sum_{\nu=0}^{\infty} \frac{r_n^\nu}{r_n^n} \frac{t_n}{t_\nu}.$$

If a sequence r_n exists for which (20) is satisfied, then we see that (19) is satisfied and hence that $\{t_n\}$ is a B-sequence.

THEOREM X. *A sufficient condition that the sequence $\{t_n\}$ of positive numbers $t_0, t_1, t_2, \cdots$ be a B-sequence is that the corresponding sequence $\{\lambda_n\}$ be monotonic non-decreasing and have infinity as a limit.*

It is clear that condition (4) is satisfied. It is easy to verify that the hypotheses imply the relations

$$\lambda_n^{\nu-n} t_n/t_\nu \leqq 1, \qquad (\nu, n = 1, 2, \cdots).$$

Hence, if we take $r_n = (1-1/n)\lambda_n$ for $n>1$, we have

$$1 \leqq \sum_{\nu=0}^{\infty} \frac{r_n^\nu}{r_n^n} \frac{t_n}{t_\nu} \leqq \sum_{\nu=0}^{\infty} \left(1-\frac{1}{n}\right)^{\nu-n} + \left(\frac{1}{t_0}-1\right)\left(1-\frac{1}{n}\right)^{-n}$$

$$= \left(n+\frac{1}{t_0}-1\right)\left(1-\frac{1}{n}\right)^{-n}, \qquad n>1.$$

Then equation (20) follows. Then Theorem X follows from Theorem IX.

THEOREM XI. *A necessary condition that the sequence* $\{t_n\}$ *of positive numbers* $t_0, t_1, t_2, \cdots$ *be a B-sequence is that we have*

$$\liminf_{n=\infty} (\lambda_{n+1}/\lambda_n)^{1/n} \geqq 1.$$

We shall prove this by showing that a contradiction arises if we assume that a *B*-sequence $\{t_n\}$ exists for which we have

$$\liminf_{n=\infty} (\lambda_{n+1}/\lambda_n)^{1/n} = \theta < 1.$$

Let θ_1 and θ_2 be positive numbers such that $\theta<\theta_1^2<\theta_1<\theta_2<1$. Then we have

$$\lambda_{n+1} < \theta_1^{2n}\lambda_n < \theta_1^n \lambda_n$$

for an infinite increasing sequence of positive values $n_0, n_1, n_2, \cdots$ of n. Then for such n we have

$$\frac{|x|^n}{t_n\theta_2^n} = \frac{|x|^{n-1}}{t_{n-1}}\left(\frac{\theta_1}{\theta_2}\right)^n \frac{|x|}{\theta_1^n\lambda_n} \leqq \frac{|x|^{n-1}}{t_{n-1}}\left(\frac{\theta_1}{\theta_2}\right)^n \leqq E(|x|)\left(\frac{\theta_1}{\theta_2}\right)^n$$

$$\text{if } |x| \leqq \theta_1^n\lambda_n,$$

$$\frac{|x|^n}{t_n\theta_2^n} = \frac{|x|^{n+1}}{t_{n+1}}\left(\frac{\theta_1}{\theta_2}\right)^n \frac{\theta_1^{-n}\lambda_{n+1}}{|x|} \leqq \frac{|x|^{n+1}}{t_{n-1}}\left(\frac{\theta_1}{\theta_2}\right)^n \leqq E(|x|)\left(\frac{\theta_1}{\theta_2}\right)^n$$

$$\text{if } |x| \geqq \theta_1^{-n}\lambda_{n+1}.$$

Since $\theta_1^{-n}\lambda_{n+1}<\theta_1^n\lambda_n$ for such n the two ranges of x cover all finite values of x. Hence, for the function

$$F(x) = \sum_{\nu=0}^{\infty} \frac{x^{n_\nu}}{t_{n_\nu}\theta_2^{n_\nu}}$$

we have the relation

$$|F(x)| \leqq F(|x|) \leqq \frac{\theta_2}{\theta_2-\theta_1} E(|x|).$$

But

$$\limsup_{n=\infty} \left| t_n F^{(n)}(0)/n! \right|^{1/n} = \frac{1}{\theta_2} > 1.$$

The last two relations contradict the hypothesis that $\{t_n\}$ is a B-sequence. Hence the theorem follows as stated.

The foregoing theorem will enable us to prove the following:

THEOREM XII. *Not every A-sequence is a B-sequence.*

Let $\{t_n\}$ denote the sequence for which we have

$$\rho_0 = 1, \qquad \rho_n = 2^\nu \text{ for } 3\cdot 2^{\nu-2} < n \leqq 3\cdot 2^{\nu-1}, \qquad (\nu = 0, 1, 2, \cdots).$$

Since ρ_n is monotonic non-decreasing and has the limit infinity it follows from Corollary 3 to Theorem II that $\{t_n\}$ is an A-sequence. But from Theorem XI it follows that it is not a B-sequence, since we now have

$$\liminf_{n=\infty} (\lambda_{n+1}/\lambda_n)^{1/n} \leqq \tfrac{1}{2},$$

as one sees by considering the values $3\cdot 2^{\nu-2}+1$ of n. Hence the theorem is established.

The sequences $\{t_n\}$ defined in equations (11) to (14) are B-sequences, as one shows by aid of Theorem X and the fact that $\lambda_n \geqq \rho_n$ when ρ_n is monotonic non-decreasing, owing to the relation $\lambda_n = \rho_n^n (\rho_{n-1})^{-n+1}$ which holds in all cases for positive n.

9. **Necessary and sufficient conditions for normal type $(t_n;\ q)$.** We have the following theorem:

THEOREM XIII. *Let $\{t_n\}$ be a B-sequence. Then a necessary and sufficient condition that a single-valued analytic function $F(x)$ shall be of normal type $(t_n; q)$ is that* (a) *numbers p shall exist for which it is true that for every positive number ϵ there exists a constant M, depending on ϵ and p but independent of x, such that for all finite values of x we have*

$$|F(x)| < ME[(p+\epsilon)|x|]$$

and (b) *that q shall be the least of such numbers p.*

Moreover, when $F(x)$ is of normal type $(t_n; q)$ we have

$$|F^{(n)}(x)| < M(q+\epsilon)^n E^{(n)}[(q+\epsilon)|x|], \qquad (n = 0, 1, 2, \cdots),$$

where M is independent of x and n.

The proposition in the first paragraph of the theorem is an immediate consequence of the earlier part of §8; that in the second paragraph follows from (7) and the fact that every B-sequence is an A-sequence.

Pringsheim (loc. cit., pp. 273, 280) has observed that the sequence $\{(n!)^{1/t}\}$, $0<t<\infty$, possesses the properties characterizing both our A-sequences and our B-sequences. Results obtained by him imply the following proposition:

A necessary and sufficient condition that a single-valued analytic function $F(x)$ shall be of normal type $\{(n!)^{1/t}; q\}$, $0<t<\infty$, $0<q<\infty$, that is to say, of the classic order t and normal type q^t/t, is that (a) numbers p shall exist for which it is true that for every positive number ϵ there exists a quantity M, depending on ϵ and p but independent of x, such that for all finite values of x we have

$$|F(x)| < Me^{(p+\epsilon)^t|x|^t/t}$$

and (b) that q shall be the least of such numbers p.

From this result and our Theorem V we have the following theorem:

THEOREM XIV. *If $F(x)$ is of normal type $\{(n!)^{1/t}; q\}$, $0<t<\infty$, then, for every positive η, ϵ, δ, there exists a constant $M_{\eta\epsilon\delta}$, independent of x and n, such that for all n and for all finite x we have the inequality*:

$$|(n!)^{1/t}F^{(n)}(x)/n!| < M_{\eta\epsilon\delta}(q+\epsilon)^n\left(1+\frac{1}{\delta}\right)^n e^{(q+\epsilon)^t(1+\delta)^t(1+\eta)^t|x|^t/t}.$$

UNIVERSITY OF ILLINOIS,
URBANA, ILL.

INTEGRATION AND LINEAR OPERATIONS*

BY

NELSON DUNFORD

1. **Notation and introduction.** By L_p $(1 \leqq p < \infty)$ will be understood the class of real valued measurable functions $\phi(P)$, $0 \leqq P \leqq 1$, for which $\int_0^1 |\phi(P)|^p dP < \infty$, by $S_p(X)$ the class of measurable† functions $f(P)$ on (0,1) to the space X of type B [see‡ 1, p. 53] for which $\int_0^1 \|f(P)\|^p dP < \infty$. With $\|f\| \equiv \{\int_0^1 \|f(P)\|^p dP\}^{1/p}$, $S_p(X)$ is a Banach space. In case $p > 1$, $L_{p'}$ is defined by the equality $p' = p/(p-1)$ and if $p = 1$ the symbol $L_{p'} = L_\infty = M$ will stand for the space of real, essentially bounded and measurable functions with $\|\phi\| = \text{ess. sup. } |\phi(P)|$. Similarly for $S_{p'}(X)$ and $S_\infty(X)$, and for brevity we write S_{pq} in place of $S_p(L_q)$. By L_{pq} we mean the class of real valued measurable functions $K(P, Q)$ that belong to L_q for each P and for which $\{\int_0^1 |K(P, Q)|^q dQ\}^{1/q}$ belongs to L_p. Finally the term *linear operation* is used in the sense of Banach, i.e., for an additive continuous function. In terms of this notation our results are described in the following paragraph.

It is easily shown that a kernel $K(P, Q)$ in $L_{p'q}$ defines a linear operation

$$T\phi = \int_0^1 K(P, Q)\phi(P)dP \tag{1}$$

on L_p to L_q. In case $p = 1$, $1 < q < \infty$, this is the expression for the most general linear transformation and its norm is

$$|T| = \text{ess. sup.} \left\{\int_0^1 |K(P, Q)|^q dQ\right\}^{1/q}.$$

In case $1 < p \leqq \infty$, $1 \leqq q < \infty$ the operation (1) is completely continuous and the general linear operation is expressible as the limit of a sequence of operations of type (1). The general completely continuous linear operation on $L = L_1$ to L_q $(1 < q < \infty)$ is given by (1), where the kernel $K(P, Q)$ is in $L_{\infty q}$, vanishes outside the square $0 \leqq P, Q \leqq 1$, and satisfies the condition

$$\lim_{h=0} \text{ess. sup.} \int_0^1 |K(P, Q + h) - K(P, Q)|^q dQ = 0.$$

---

* Presented to the Society, October 26, 1935, and February 29, 1936; received by the editors, September 1, 1935, and, in revised form, February 11, 1936.

† This concept is defined in §2 of this paper.

‡ References in brackets refer to the bibliography at the end of the paper.

Other representation theorems are given for operations on L to l_p (the space of sequences $\{a_i\}$ of real numbers with $\sum_{i=1}^{\infty}|a_i|^p$ convergent), and operations on the space of absolutely continuous functions to L_p, l_p, or Hilbert space. The abstract function $f(P)=K(P, Q)$ may be in S_{pq} without the real function $K(P, Q)$ being in L_{pq} but every function in S_{pq} has a unique representation in L_{pq} and conversely every function $K(P, Q)$ in L_{pq} defines uniquely a function $f(P)=K(P, Q)$ in S_{pq}.

2. **Preliminary remarks.** A function $f(P)$ on $0\leqq P\leqq 1$ to a Banach space X is said to be *finitely valued** in case there exists a decomposition of the interval (0,1) into a finite number of disjoint measurable sets on each of which $f(P)$ is constant, and it is said to be *measurable* in case there exists a sequence $f_n(P)$ of finitely valued functions for which $f(P)=\lim_n f_n(P)$ almost everywhere. This definition of measurability which was used by Bochner is equivalent to saying that $f(P)$ satisfies the condition of Lusin, i.e., $f(P)$ is continuous on a closed set with measure arbitrarily near one. This fact is proved almost as in the case of real functions [13, p. 44].

S. Bochner has defined [3] a class of functions called summable on $(0, 1)$ to X and an integral on this class. We [6] have done likewise and in both cases the space of summable functions is complete and has the set of finitely valued functions everywhere dense. This shows that the two notions of summable functions as well as the integrals coincide. Accordingly results from both papers will be used. The reader will find no difficulty in establishing the following theorems (the proofs are merely sketched) which are stated here and numbered for future reference.

2.1. *If $f(P)$ is measurable and $\phi(P)$ is a real finite valued measurable function then the functions $\|f(P)\|$ and $\phi(P)f(P)$ are measurable. $f(P)/\phi(P)$ is measurable if an arbitrary constant value is assigned to it on the set where $\phi(P)=0$. Also if a sequence of measurable functions $f_n(P)$ converges almost everywhere then $\lim_n f_n(P)$ is measurable.*

2.2. *The function $f(P)$ is summable if and only if $f(P)$ is measurable and $\|f(P)\|$ is summable.*

This is Bochner's definition of a summable function.

2.3. *Let X and Y be two spaces of type B, Tx linear on X to Y and $f(P)$ summable on $(0, 1)$ to X, then $Tf(P)$ is summable and $T\int_e f(P)dP=\int_e Tf(P)dP$ for every measurable subset e of $(0, 1)$.*

That $Tf(P)$ is summable follows from 2.2. For $Tf(P)$ is continuous on

* The reader should be careful to distinguish between finitely valued and finite valued, the latter term being used in the usual sense for real functions.

any set upon which $f(P)$ is continuous so that $Tf(P)$ is measurable. The norm $\|Tf(P)\|$ is summable since it is bounded by $|T| \cdot \|f(P)\|$. The equality in the theorem is proved by using a sequence of finitely valued functions converging to $f(P)$ in the mean. This theorem has been given by G. Birkhoff [2].

2.4. *If $f(P)$ is in $S_{p'}(X)$ then the transformation*

$$T\phi = \int_0^1 f(P)\phi(P)dP$$

is a linear operation on L_p to X, with $|T| \leqq \{\int_0^1 \|f(P)\|^{p'}dP\}^{1/p'}$ if $1 < p \leqq \infty$ and $|T| \leqq$ ess. sup. $\|f(P)\|$ if $p = 1$.

The operation is defined on L_p, for by 2.1, $f(P)\phi(P)$ is measurable and since $\|f(P)\phi(P)\| = \|f(P)\| \cdot |\phi(P)|$ is summable it follows from 2.2 that $f(P)\phi(P)$ is summable. $T\phi$ is linear since $\|T\phi\| \leqq \int_0^1 \|f(P)\| \cdot |\phi(P)|\, dP \leqq \|f\| \cdot \|\phi\|$.

3. **The representation of summable and of measurable functions.** A few preliminary remarks are in order. It is easily seen that if the abstract function $f(P) = K(P, Q)$ is in S_{pq}, $1 \leqq p, q < \infty$, then

$$\int_0^1 dP \left\{ \int_0^1 |K(P, Q)|^q dQ \right\}^{p/q} = \int_0^1 \|f(P)\|^p dP < \infty,$$

but this does not mean that $K(P, Q)$ is in L_{pq} for $K(P, Q)$ may not be measurable. In fact it may be that $f(P) = 0$ for all P and that $K(P, Q)$ is a non-measurable function. This is the case if $K(P, Q)$ is the characteristic function of a Sierpiński set [15]. Such a set is bounded and has positive exterior measure and has no more than two points in common with any straight line. Also one might ask if every function $K(P, Q)$ for which

$$\int_0^1 dP \left\{ \int_0^1 |K(P, Q)|^q dQ \right\}^{p/q} < \infty$$

defines a function $f(P) = K(P, Q)$ in S_{pq}. That such is not the case is seen by the example $K(P, Q) = \phi(P)\psi(Q)$, where $\psi(Q)$ is any function in L_q and $\phi(P)$ is a non-measurable function with $|\phi(P)|$ in L_p. However, as will be seen, every $K(P, Q)$ in L_{pq} defines a function in S_{pq} and conversely every $f(P)$ in S_{pq} has a representation in L_{pq} and two such representations must (by the theorem of Fubini) be equal for all P, Q except for a set of two-dimensional measure zero. In fact for a large class of Banach spaces Y a measurable function $f(P)$ on (0,1) to Y has a measurable representation. For L_2, this fact as well as 3.2 has been established by S. Bochner and J. von Neumann [4]. Let Y be a Banach space composed of real almost everywhere finite measurable

functions $\psi=\psi(Q)$ defined for $0\leqq Q\leqq 1$. Let it be supposed that Y satisfies the following conditions:

(a) If $\|\psi_n\|\to 0$ then $\psi_n(Q)\to 0$ in measure.

(b) If $\psi(Q)=0$ almost everywhere then $\psi=0$.

Then we have

3.1. *If the function $f(P)$ on $(0,1)$ to Y is measurable, there exists a function $K(P, Q)$ measurable on the square $0\leqq P, Q\leqq 1$ such that $f(P)=K(P, Q)$. Any two measurable representations of $f(P)$ differ only on a set of two-dimensional measure zero.*

First suppose that $f(P)=K^*(P, Q)$ is finitely valued, and let $e_1, e_2, \cdots, e_n$ be the sets upon which $f(P)$ is constant. Let P_i be a point of e_i and define

$$K(P, Q) = K^*(P_i, Q) \text{ for } P \text{ in } e_i \qquad (i = 1, 2, \cdots, n).$$

So that $K(P, Q)$ is a measurable representation of $f(P)$. Now suppose $f(P)=K^*(P, Q)$ is an arbitrary measurable function, and let $f_n(P)$ be a sequence of finitely valued functions approaching $f(P)$ almost everywhere. Let $K_n(P, Q)$ be a measurable representation of $f_n(P)$, then

$$\lim_{n,m} \|K_n(P,Q) - K_m(P,Q)\|_Y = \lim_{n,m} \|f_n(P) - f_m(P)\| = 0 \text{ for almost all } P.$$

Thus by (a), $K_n(P, Q)$ converges in measure for almost all P in $(0, 1)$ and so we have

$$\lim_{n,m} \int_0^1 \frac{|K_n(P, Q) - K_m(P, Q)|}{1 + |K_n(P, Q) - K_m(P, Q)|}\, dQ = 0 \text{ for almost all } P,$$

and hence (using the theorem of Fubini)

$$\lim_{n,m} \int_0^1 \int_0^1 \frac{|K_n(P, Q) - K_m(P, Q)|}{1 + |K_n(P, Q) - K_m(P, Q)|}\, dP dQ = 0.$$

There is therefore a function $K'(P, Q)$ measurable on the square $0\leqq P, Q\leqq 1$ such that

$$K_m(P, Q) \to K'(P, Q) \text{ in measure on } 0 \leqq P, Q \leqq 1.$$

Hence there is a subsequence $K_{m_i}(P, Q)$ and a set E_1 in $(0, 1)$ with $m(E_1)=1$ such that

$$\lim_{i=\infty} \int_0^1 \frac{|K_{m_i}(P, Q) - K'(P, Q)|}{1 + |K_{m_i}(P, Q) - K'(P, Q)|}\, dQ = 0 \text{ for } P \text{ in } E_1,$$

that is, for P in E_1

$$K_{m_i}(P, Q) \to K'(P, Q) \text{ in measure.}$$

Now since $\|f_n(P) - f(P)\| \to 0$ for almost all P, say for P in E_2 where $m(E_2) = 1$, we have (by (a)) for P in E_2

$$K_m(P, Q) \to K^*(P, Q) \text{ in measure.}$$

Thus for P in E_1E_2

$$K'(P, Q) = K^*(P, Q) \text{ for almost all } Q,$$

and so by (b) $f(P) = K'(P, Q)$ for P in E_1E_2. By defining

$$\begin{aligned} K(P, Q) &= K'(P, Q) \text{ for } P \text{ in } E_1E_2, \\ &= K^*(P, Q) \text{ elsewhere,} \end{aligned}$$

we have $f(P) = K(P, Q)$ for $0 \leqq P \leqq 1$. Hence since K differs from K' on a set of two-dimensional measure zero it is a measurable representation of $f(P)$.

Now suppose $K(P, Q)$ and $K'(P, Q)$ are two different measurable representations of $f(P)$. From (a) it follows that $\psi = 0$ implies $\psi(Q) = 0$ almost everywhere and hence for all P

$$K(P, Q) = K'(P, Q) \text{ for almost all } Q,$$

so that by the theorem of Fubini $K(P, Q) = K'(P, Q)$ almost everywhere on $0 \leqq P, Q \leqq 1$.

The converse of the theorem is not true in general; i.e., if $K(P, Q)$ is a measurable function of the two variables P and Q, and if for each P, $K(P, Q)$ is in Y then it does not follow that the abstract function $f(P) = K(P, Q)$ on $(0, 1)$ to Y is measurable. For take $Y = M$ the space of real, essentially bounded and measurable functions defined on $(0, 1)$, and define

$$\begin{aligned} K(P, Q) &= 0, \qquad P \leqq Q, \\ &= 1, \qquad P > Q. \end{aligned}$$

Then the abstract function $f(P) = K(P, Q)$ has the property that for $P < P'$, $\|f(P) - f(P')\| = 1$ and consequently does not satisfy the condition of Lusin. As we shall see in 3.2 this phenomenon does not occur if M is replaced by L_p.

Hereafter when we speak of a function $f(P)$ in S_{pq} and write $f(P) = K(P, Q)$ it will be understood, without explicit mention of the fact that $K(P, Q)$ is measurable. It should be noted that if $K(P, Q)$ is in L_{pq} then for every finite valued $\phi(P)$ in $L_{p'}$, $K(P, Q)\phi(P)$ is in L_{1q} and is a measurable representation of $f(P)\phi(P)$.

3.2. *The function $f(P)$ is in S_{pq}, $1 \leqq p, q < \infty$, if and only if there exists a function $K(P, Q)$ in L_{pq} such that $f(P) = K(P, Q)$. If $f(P)$ is in S_{pq} then*

$$\int_e f(P)dP = \int_e K(P, Q)dP$$

for every measurable subset e of (0, 1).

First suppose that $K(P, Q)$ is in L_{pq}, then

$$\|f(P)\|^p = \left\{ \int_0^1 | K(P, Q) |^q dQ \right\}^{p/q}$$

is summable and it only remains to show that $f(P)$ is measurable. Let $\psi_n(Q)$ be the orthonormal set of Haar [8]. Then for each P

$$K(P, Q) = f(P) = \sum_{n=1}^{\infty} \phi_n(P)\psi_n(Q),$$

where

$$\phi_n(P) = \int_0^1 K(P, Q)\psi_n(Q)dQ.$$

This means that $\sum_1^n \phi_n(P)\psi_n(Q)$ converges in the norm of L_q to $K(P, Q)$ [14]. Since $K(P, Q)$ is measurable so is $\phi_n(P)$ and thus by 2.1 the abstract function

$$f_n(P) = \sum_1^n \phi_n(P)\psi_n$$

is measurable. Since $f_n(P) \to f(P)$ it follows again by 2.1 that $f(P)$ is measurable and thus $f(P)$ is in S_{pq}.

Now conversely suppose $f(P)$ is in S_{pq}. By 3.1 there is a measurable function $K(P, Q)$ such that $f(P) = K(P, Q)$ and $K(P, Q)$ is obviously in L_{pq}. To establish the equality stated in the theorem we again use the Haar base and recall [1, p. 111] that for each x in L_q

$$x = \sum_1^{\infty} \psi_i T_i x,$$

where T_i is a linear function on L_q. If we write

$$\phi_i(P) = T_i f(P), \qquad |\phi_i(P)| \leqq |T_i| \cdot \|f(P)\|$$

it is seen that $\phi_i(P)$ is a finite valued function in L_p and

$$K(P, Q) = f(P) = \sum_1^{\infty} \phi_i(P)\psi_i(Q).$$

From 2.3 it follows that

$$\int_e f(P)dP = \sum_1^\infty \psi_i T_i \int_e f(P)dP = \sum_1^\infty \psi_i \int_e T_i f(P)dP = \sum_1^\infty \psi_i \int_e \phi_i(P)dP,$$

which means that

$$\xi_n(Q) \equiv \sum_1^n \psi_i(Q) \int_e \phi_i(P)dP$$

approaches $\int_e f(P)dP$ in L_q as $n \to \infty$. To complete the proof it is only necessary to show that this same sequence approaches $\int_e K(P, Q)dP$ in L_q. Since $\xi_n(Q)$ converges in L_q it is sufficient to show that $\xi_n(Q) \to \int_e K(P, Q)dP$ in measure. Now writing $K_n(P, Q) = \sum_1^n \phi_i(P)\psi_i(Q)$, we have for each P [14]

$$\left\{ \int_0^1 | K_n(P, Q) - K(P, Q) |^q dQ \right\}^{1/q} \to 0$$

and

$$\begin{aligned} \left\{ \int_0^1 | K_n(P, Q) - K(P, Q) |^q dQ \right\}^{1/q} &\leqq \left\{ \int_0^1 | K_n(P, Q) |^q dQ \right\}^{1/q} \\ &\quad + \left\{ \int_0^1 | K(P, Q) |^q dQ \right\}^{1/q} \\ &\leqq 2 \left\{ \int_0^1 | K(P, Q) |^q dQ \right\}^{1/q}. \end{aligned}$$

Thus the sequence on the left side of this inequality is bounded by a function in L_p, and so

$$\begin{aligned} &\int_0^1 dQ \int_0^1 | K_n(P, Q) - K(P, Q) | dP \\ &\qquad \leqq \left(\int_0^1 dP \left\{ \int_0^1 | K_n(P, Q) - K(P, Q) |^q dQ \right\}^{p/q} \right)^{1/p} \to 0, \end{aligned}$$

from which the desired conclusion is easily deduced. This completes the proof of 3.2.

3.3. *The function $f(P)$ on $(0, 1)$ to L_q, $1 \leqq q < \infty$, is measurable and essentially bounded if and only if there exists a measurable function $K(P, Q)$ in L_q for each P and satisfying the conditions*

$$\text{ess. sup.} \int_0^1 | K(P, Q) |^q dQ < \infty,$$

$$f(P) = K(P, Q),$$

and in case the conditions are satisfied

$$\int_e f(P)dP = \int_e K(P, Q)dP$$

for every measurable subset e of (0, 1).

This is an immediate corollary of 3.2.

3.4. *The function* $f(P) = \{k_i(P)\}$ *is in* $S_p(l_q)$ $(1 \leqq p \leqq \infty, 1 \leqq q < \infty)$ *if and only if*

(i) $k_i(P)$ *is measurable* $(i = 1, 2, 3, \cdots)$;

(ii) $\left\{ \sum_1^\infty | k_i(P) |^q \right\}^{1/q}$ *is in* L_p.

For functions $f(P) = \{k_i(P)\}$ *in* $S_p(l_q)$ *we have*

$$\int_e f(P)dP = \left\{ \int_e k_i(P)dP \right\}$$

for every measurable subset e of (0, 1).

This theorem can be established by a method quite analogous but much simpler than the method used in the proof of 3.2. In fact if we take $\psi_i = \{\delta_k^i\}$ (the Kronecker delta) as the base for l_q, then

(iii) $$f(P) = \sum_1^\infty T_i f(P)\psi_i = \sum_1^\infty k_i(P)\psi_i,$$

which shows that $k_i(P)$ is measurable providing $f(P)$ is. Condition (ii) is merely a statement of the fact that $\|f(P)\|$ is in L_p. Conversely if (i) and (ii) are satisfied, (iii) combined with 2.1 shows that $f(P)$ is measurable and (ii) shows that $\|f(P)\|$ is in L_p. The equality $\int_e f(P)dP = \{\int_e k_i(P)dP\}$ is merely an application of 2.3, for we have

$$\int_e f(P)dP = \sum_1^\infty \psi_i T_i \int_e f(P)dP = \sum_1^\infty \psi_i \int_e T_i f(P)dP$$

$$= \sum_1^\infty \psi_i \int_e k_i(P)dP = \left\{ \int_e k_i(P)dP \right\}.$$

4. **Applications of the preceding results.** In view of 3.2 each theorem regarding functions in S_{pq} may be translated into a theorem concerning functions in L_{pq}. Since it is merely a translation of symbols, we shall content ourselves here with one illustration. For example, Theorem 9 of reference 6 reads as follows:

4.1. *If the functions* $K(P, Q)$, $K_m(P, Q)$, $m=1, 2, 3, \cdots$, *are in* L_{pq} $(1 \leqq p, q < \infty)$, *and if* $\int_0^1 |K_m(P, Q) - K(P, Q)|^q dQ$ *approaches zero in measure, then the following assertions are equivalent*:

(1) $$\lim_m \int_0^1 dQ \left| \int_e \left\{ K_m(P,Q) - K(P,Q) \right\} dP \right|^q = 0$$

for every measurable set e *in* $(0, 1)$.

(2) $$\lim_m \int_0^1 dQ \left| \int_e \left\{ K_m(P,Q) - K(P,Q) \right\} dP \right|^q = 0$$ *uniformly for* e *in* $(0, 1)$.

(3) $$\lim_{n,m} \int_0^1 dQ \left| \int_e \left\{ K_m(P,Q) - K_n(P,Q) \right\} dP \right|^q = 0$$ *for each* e *in* $(0, 1)$.

(4) $$\lim_{m(e)=0} \limsup_m \int_0^1 dQ \left| \int_e K_m(P,Q) dP \right|^q = 0.$$

(5) $$\lim_{m(e)=0} \int_0^1 dQ \left| \int_e K_m(P, Q) dP \right|^q = 0$$ *uniformly with respect to* m.

5. **The representation of linear operations.** We have

5.1. *If* X *is of type B with a base, then* $T\phi$ *is a linear operation on* L_p $(1 \leqq p \leqq \infty)$ *to* X *if and only if there exists a sequence of functions* $f_n(P)$ *in* $S_{p'}(X)$ *such that* $\int_0^1 f_n(P)\phi(P)dP$ *converges for each* $\phi(P)$ *in* L_p *and*

$$T\phi = \lim_n \int_0^1 f_n(P)\phi(P)dP \text{ for } \phi(P) \text{ in } L_p.$$

Let T be a linear operation on L_p to X, and let $\{x_i\}$ be a base for X, then

$$T\phi = \sum_1^\infty a_i(\phi)x_i,$$

where $a_i(\phi) = T_i T\phi$ is a linear functional on L_p. Thus there is a finite valued function $\alpha_i(P)$ in $L_{p'}$ such that [1, p. 64]

$$a_i(\phi) = \int_0^1 \alpha_i(P)\phi(P)dP.$$

Since $\alpha_i(P)x_i$ is in $S_{p'}(X)$, $f_n(P) = \sum_1^n \alpha_i(P)x_i$ is likewise and the sequence

$$\int_0^1 f_n(P)\phi(P)dP = \sum_1^n a_i(\phi)x_i$$

converges for each ϕ in L_p to $T\phi$. The converse statement follows from 2.4 and a well known result [1, p. 23, Theorem 4].

5.2. *The function $T\phi$ is a linear operation on $L_p(1 \leqq p \leqq \infty)$ to $L_q(1 \leqq q < \infty)$ if and only if there exists a sequence of measurable functions $K_n(P, Q)$ belonging to L_q for each P, and satisfying the conditions*

$$\left\{\int_0^1 |K_n(P, Q)|^q dQ\right\}^{1/q} \text{ is in } L_{p'} \text{ for each } n,$$

and

$$T\phi = \lim_n \int_0^1 K_n(P, Q)\phi(P)dP \text{ for } \phi \text{ in } L_p.$$

This follows immediately from 5.1, 3.2, and 3.3.

In the case of transformations $T\phi$ on L to X, where X is a Banach space satisfying one of the conditions:

(A) X has a uniformly convex norm [5];

(B) X has a base $\{x_i\}$ such that $\sum_1^\infty a_i x_i$ converges whenever $\{a_i\}$ is a sequence of constants for which $\|\sum_1^n a_i x_i\|$ is bounded,

further results may be given. The spaces satisfying one or the other of these conditions include Hilbert space, $l_p(1 \leqq p < \infty)$, and $L_p(1 < p < \infty)$. The fundamental fact [5, 7] in this connection is that a function $F(P)$ on (0, 1) to such a space, which satisfies a Lipschitz condition $\|F(P) - F(P')\| \leqq M|P - P'|$ (or even a more general condition) has a derivative $F'(P)$ almost everywhere which is summable and such that

$$F(P) = \int_0^P F'(P)dP + F(0).$$

We shall first give a generalization of this result which is embodied in 5.3. The corresponding theorem for real functions has been given for special cases by J. Radon and P. J. Daniell and in the general case by O. Nikodym [13, pp. 255–257]. We shall use the terminology of Saks [13, pp. 247–257] and denote by $m(e)$ a completely additive non-negative function defined on an additive family A of sets of abstract elements P belonging to a space E. The sets of A are called measurable sets and the number $m(e)$ is called the measure of the set e. It is assumed that all subsets of a set of measure zero are measurable. An additive set function $F(e)$ on A to a Banach space is said to be of bounded variation in case $\sum_{i=1}^n \|F(e_i)\|$ is bounded with respect to all choices of a finite number of disjoint measurable sets $e_1, e_2, \cdots, e_n$. The l.u.b. of $\sum_{i=1}^n \|F(e_i)\|$, where $e_1, e_2, \cdots, e_n$ are disjoint measurable subsets of a measurable set e is called the total variation of F on e and is denoted by $T(F, e)$.

5.3. *Every completely additive set function $F(e)$ defined on A, which is of bounded variation and vanishes on sets of measure zero and whose values lie in a Banach space X which satisfies condition* (B) *is representable in the form*

$$F(e) = \int_e f(P)dm.$$

The function $f(P)$ is uniquely determined except for a set of measure zero.

It is no loss of generality to assume [7]

(C) $$\left\| \sum_{i=1}^{n} a_i x_i \right\| \leqq \left\| \sum_{i=1}^{n+1} a_i x_i \right\|.$$

Now suppose

$$F(e) = \sum_{i=1}^{\infty} a_i(e)x_i = \sum_{i=1}^{\infty} x_i \int_e a_i(P)dm,$$

where in the last equality we have used the theorem of Nikodym. Set

$$F_n(e) = \sum_{i=1}^{n} a_i(e)x_i, \qquad f_n(P) = \sum_{i=1}^{n} a_i(P)x_i.$$

By a well known computation and the inequality (C) we get

$$T(F_n, e) = \int_e \|f_n(P)\|dm \leqq T(F, E)$$

$$\|f_n(P)\| \leqq \|f_{n+1}(P)\|.$$

Thus if $b(P) \equiv \lim_n \|f_n(P)\|$ it is seen that $b(P)$ is summable. Hence by (B)

$$f(P) = \sum_{i=1}^{\infty} a_i(P)x_i$$

is convergent for almost all P and $\|f(P)\| = b(P)$, so that the function $f(P)$ is summable and

$$\int_e f(P)dm = \sum_{i=1}^{\infty} x_i \int_e a_i(P)dm = \sum_{i=1}^{\infty} a_i(e)x_i = F(e).$$

Suppose we have also $F(e) = \int_e f^*(P)dm$, where $f^*(P) = \sum_{i=1}^{\infty} a_i^*(P)x_i$. Then $\int_e a_i^*(P)dm = \int_e a_i(P)dm$ for each e and so $a_i^*(P) = a_i(P)$ almost everywhere and thus the same is true of $f^*(P)$ and $f(P)$.

In the case of an absolutely continuous additive function $F(R)$ of bounded variation on elementary figures in a fundamental interval of Euclidean n-space to a Banach space with a uniformly convex norm Clarkson [5] has shown that

$$F(R) = \int_R F'(P)dP,$$

where R is an elementary figure. Hence in case the Banach space X satisfies either condition (A) or (B) we can say that an additive function $F(e)$ on Lebesgue measurable sets in (0, 1) to X which satisfies the Lipschitz condition $\|F(e)\| \leqq M|e|$ also satisfies the conditions

$$F(e) = \int_e F'(P)dP, \qquad \|F'(P)\| \leqq M.$$

Such a function $F(e)$ is defined by a linear operator T on L to X by placing $F(e) = T\phi(e)$, where $\phi(e)$ is the characteristic function of the set e. Here $M = |T|$. Thus for all finitely valued functions $\phi(P)$

$$T\phi = \int_0^1 F'(P)\phi(P)dP.$$

By taking a sequence of finitely valued functions converging in the mean to an arbitrary summable function it is readily seen that the above equality holds for all ϕ in L and that

$$|T| \leqq \sup\|F'(P)\| \leqq M = |T|.$$

Thus we have

5.4. *The function $T\phi$ is a linear operation on L to X (where X satisfies* (A) *or* (B)) *if and only if there exists an essentially bounded and measurable function $f(P)$ on* (0, 1) *to X such that*

$$T\phi = \int_0^1 f(P)\phi(P)dP.$$

The norm of T is $|T| = \text{ess. sup. } \|f(P)\|$.

As a corollary we have (using 3.2)

5.5. *The function $T\phi$ is a linear operation on L to L_q $(1 < q < \infty)$ if and only if there exists a measurable function $K(P, Q)$ in L_q for each P and satisfying the conditions*

$$\text{ess. sup.} \int_0^1 |K(P, Q)|^q dQ < \infty,$$

$$T\phi = \int_0^1 K(P, Q)\phi(P)dP \text{ for } \phi \text{ in } L.$$

The norm of T is $|T| = \text{ess. sup. } \{\int_0^1 |K(P, Q)|^q dQ\}^{1/q}$.

As would be expected, the method of constructing the absolutely con-

tinuous function $F(P)=T\phi(0, P)$ from the operation T does not work in case the domain of T is L_p, $p>1$, instead of L. In this case

$$\sum_1^n \|F(b_i)-F(a_i)\| \leqq |T| \sum_1^n |b_i-a_i|^{1/p}.$$

But now it is not in general true that

$$\sum_1^n |b_i-a_i|^{1/p} \to 0 \quad \text{as} \quad \sum_1^n |b_i-a_i| \to 0,$$

for suppose ϵ is any positive number, then for $n>(2^p\epsilon)^{1/(1-p)}$ and $|b_i-a_i| = \epsilon/n$ we have $\sum_1^n |b_i-a_i| = \epsilon$ and $\sum_1^n |b_i-a_i|^{1/p}>1/2$.

Using 3.4 we have

5.6. *The function $T\phi$ is a linear operation on L to l_q $(1\leqq q<\infty)$ if and only if there exists a sequence $k_i(P)$ of measurable functions such that*

$$\text{ess. sup.} \sum_1^\infty |k_i(P)|^q < \infty,$$

$$T\phi = \left\{ \int_0^1 k_i(P)\phi(P)dP \right\}.$$

The norm of T is $|T|=$ ess. sup. $\{\sum_1^\infty |k_i(P)|^q\}^{1/q}$.

The space $A.C.$ of real functions absolutely continuous on (0, 1) which vanish at the origin, with the norming operation $\|\phi\| = V(\phi) =$ total variation of $\phi(P)$ on (0, 1) is a space of type B. Furthermore the linear operation $d\phi/dP=\phi'$ establishes the fact that $A.C.$ and L are isometric, isomorphic, and thus equivalent in the sense of Banach [1, p. 180]. The following theorems are therefore corollaries of 5.4, 5.5, and 5.6.

5.7. *The function $T\phi$ is a linear operation on $A.C.$ to the space X (X satisfying* (A) *or* (B)) *if and only if there exists an essentially bounded and measurable function $f(P)$ on* (0, 1) *to X such that*

$$T\phi = \int_0^1 f(P)\phi'(P)dP.$$

If T is linear then $|T|=$ ess. sup. $\|f(P)\|$.

5.8. *The function $T\phi$ is a linear operation on $A.C.$ to L_q $(1<q<\infty)$ if and only if there exists a measurable function $K(P, Q)$ in L_q for each P and satisfying the conditions*

$$\text{ess. sup.} \int_0^1 |K(P, Q)|^q dQ < \infty,$$

$$T\phi = \int_0^1 K(P, Q)\phi'(P)dP \text{ for } \phi \text{ in } A.C.$$

If T is linear then $|T| = \text{ess. sup.} \{\int_0^1 |K(P, Q)|^q dQ\}^{1/q}$.

5.9. *The function $T\phi$ is a linear operation on A.C. to l_q $(1 \leqq q < \infty)$ if and only if there exists a sequence $k_i(P)$ of measurable functions such that*

$$\text{ess. sup.} \sum_1^\infty |k_i(P)|^q < \infty,$$

$$T\phi = \int_0^1 k_i(P)\phi'(P)dP \text{ for } \phi \text{ in } A.C.$$

The norm of T is $|T| = \text{ess. sup.} \{\sum_1^\infty |k_i(P)|^q\}^{1/q}$.

6. Completely continuous transformations and further properties of the general integral. In the last part of the proof of 3.2, put

$$T_n\phi = \int_0^1 K_n(P, Q)\phi(P)dP, \qquad T\phi = \int_0^1 K(P, Q)\phi(P)dP.$$

It is seen that

$$|T_n - T| \leqq \left(\int_0^1 dP \left\{ \int_0^1 |K_n(P, Q) - K(P, Q)|^q dQ \right\}^{p/q} \right)^{1/p} \to 0,$$

and thus since T_n is a completely continuous linear operation on $L_{p'}$ to L_q, it follows from a well known result [1, p. 96, Theorem 2] that T is also. This fact has been established by E. Hille and J. D. Tamarkin [10] and by T. H. Hildebrandt.* Writing p' in place of p we may state

6.1. *For $1 < p \leqq \infty$, $1 \leqq q < \infty$ the transformation on L_p to L_q defined by*

$$T\phi = \int_0^1 K(P, Q)\phi(P)dP$$

with $K(P, Q)$ in $L_{p'q}$ is completely continuous.

Combined with 3.1 this gives

6.2. *For $1 < p \leqq \infty$, $1 \leqq q < \infty$ the function $f(P)$ in $S_{p'q}$ defines a completely continuous transformation*

* The proof given here is due to Hildebrandt.

$$T\phi = \int_0^1 f(P)\phi(P)dP$$

on L_p to L_q.

This result gives the integral that we are using more properties in common with the Lebesgue integral of numerical functions. The ones we have in mind center around the fact that a completely continuous linear transformation takes weakly convergent sequences into convergent sequences* [1, p. 143].

6.3. *Let $\phi_n(P)$, $n=0, 1, 2, \cdots$, be real functions and $1<p<\infty$, $1\leqq q<\infty$, then for every $f(P)$ in $S_{p'q}$ the integrals $\int_0^1 f(P)\phi_n(P)dP$ are defined and approach $\int_0^1 f(P)\phi_0(P)dP$ if and only if $\phi_n(P)$, $n=0, 1, 2, \cdots$, is in L_p and*

$$\int_0^1 |\phi_n(P)|^p dP \leqq M, \qquad n = 0, 1, 2, \cdots, \text{ for some } M,$$

$$\int_0^P \phi_n(P)dP \to \int_0^P \phi_0(P)dP, \qquad 0 \leqq P \leqq 1.$$

That the last conditions imply the convergence of the integrals follows from the statement preceding Theorem 6.2 and a theorem of F. Riesz [1, p. 135]. Conversely, if $f(P)\phi_n(P)$ is summable for every $f(P)$ in $S_{p'q}$ it is seen first, by taking $f(P)=k(P, Q)=1$ $(0\leqq P, Q\leqq 1)$ and using 3.2, that there is a summable function $K_n(P, Q)$ such that for every P

$$K_n(P, Q) = \phi_n(P) \text{ for almost all } Q.$$

It follows from the theorem of Fubini that $\phi_n(P)$ is measurable. Now let $\psi(P)$ be an arbitrary finite valued function in $L_{p'}$, and define $f(P)$ to be the function on $(0, 1)$ to L_q that takes each point P into the function in L_q which is constant and equal to $\psi(P)$, i.e., $f(P)=\psi(P)$. Then by 3.1, $\psi(P)\phi_n(P)$ is summable and

$$\int_0^1 f(P)\phi_n(P)dP = \int_0^1 \psi(P)\phi_n(P)dP.$$

From the summability of $\psi(P)\phi_n(P)$ it follows that the function $\phi_n(P)$ is in L_p, and the remaining conditions on the sequence follow from the convergence of the sequence $\int_0^1 f(P)\phi_n(P)dP$ and the theorem of Riesz. In case $p=\infty$ we have

* It is easily seen that this property characterizes completely continuous linear transformations between arbitrary Banach spaces X and Y in case the domain X is weakly complete and the conjugate space $\overline{X}$ is separable, or in case every bounded set in X is weakly compact.

6.4. *Let $\phi_n(P)$ $(n=0, 1, \cdots)$ be real functions and $1 \leqq q < \infty$, then for every summable function $f(P)$ on $(0, 1)$ to L_q the integrals $\int_0^1 f(P)\phi_n(P)dP$ are defined and approach $\int_0^1 f(P)\phi_0(P)dP$ if and only if $\phi_n(P)$ are measurable,*

$$\int_0^P \phi_n(P)dP \to \int_0^P \phi_0(P)dP, \qquad 0 \leqq P \leqq 1,$$

$$\text{ess. } \sup_P |\phi_n(P)| \leqq M, \qquad n = 0, 1, \cdots, \text{ for some } M.$$

Suppose the last conditions are satisfied and $f_0(P)$ is finitely valued with

$$\int_0^1 \|f_0(P) - f(P)\|dP < \epsilon.$$

Then by 6.3

$$\int_0^1 f_0(P)\phi_n(P)dP \to \int_0^1 f_0(P)\phi_0(P)dP,$$

and since

$$\left\| \int_0^1 f(P)(\phi_n(P) - \phi_0(P))dP \right\| \leqq \left\| \int_0^1 (f(P) - f_0(P))(\phi_n(P) - \phi_0(P))dP \right\| + \left\| \int_0^1 f_0(P)(\phi_n(P) - \phi_0(P))dP \right\|,$$

it follows that

$$\overline{\lim_n} \left\| \int_0^1 f(P)\phi_n(P)dP - \int_0^1 f(P)\phi_0(P)dP \right\| \leqq 2M\epsilon.$$

Now conversely if the integrals $\int_0^1 f(P)\phi_n(P)dP$ are defined for every summable $f(P)$ it follows as before that the functions $\phi_n(P)$ are measurable and that each $\phi_n(P)$ is essentially bounded. Also as in 6.3 if we let $\psi(P)$ be an arbitrary finite valued summable function and define $f(P)=\psi(P)$ and consider $f(P)$ as a function on $(0, 1)$ to L_q then

$$\int_0^1 f(P)\phi_n(P)dP = \int_0^1 \psi(P)\phi_n(P)dP.$$

Now the boundedness of the sequence $\int_0^1 \psi(P)\phi_n(P)dP$ for every function $\psi(P)$ in L insures the uniform boundedness of ess. $\sup_P |\phi_n(P)|$, and the convergence of the sequence $\int_0^1 \psi(P)\phi_n(P)dP$ insures by the theorem of Riesz the condition

$$\int_0^P \phi_n(P)dP \to \int_0^P \phi_0(P)dP.$$

Theorem 6.4 may be regarded as a generalization of the Riemann-Lebesgue theorem.

6.5. *The function $T\phi$ is a completely continuous linear operation on L to L_q $(1<q<\infty)$ if and only if there exists a measurable kernel $K(P, Q)$ which vanishes outside the unit square and satisfies the conditions*

$$\text{ess. sup.} \int_0^1 |K(P, Q)|^q dQ < \infty,$$

$$\lim_{h=0} \text{ess. sup.} \int_0^1 |K(P, Q+h) - K(P, Q)|^q dQ = 0,$$

$$T\phi = \int_0^1 K(P, Q)\phi(P)dP \text{ for } \phi(P) \text{ in } L.$$

The norm of T is $|T| = \text{ess. sup.} \{\int_0^1 |K(P, Q)|^q dQ\}^{1/q}$.

If the conditions are satisfied, 5.5. shows that $T\phi$ is linear. Now let S be the unit sphere in L and

$$\psi(Q) = \int_0^1 K(P, Q)\phi(P)dP.$$

It is only necessary to show that

$$\lim_{h=0} \int_0^1 |\psi(Q+h) - \psi(Q)|^q dQ = 0$$

uniformly with respect to ψ in TS [see 10, p. 445, or 12]. We have for ϕ in S

$$\left(\int_0^1 |\psi(Q+h) - \psi(Q)|^q dQ\right)^{1/q}$$

$$= \left(\int_0^1 \left|\int_0^1 \{K(P, Q+h) - K(P, Q)\}\phi(P)dP\right|^q dQ\right)^{1/q}$$

$$\leqq \int_0^1 dP\,|\phi(P)| \left\{\int_0^1 |K(P, Q+h) - K(P, Q)|^q dQ\right\}^{1/q}$$

$$\leqq \text{ess. sup.} \left\{\int_0^1 |K(P, Q+h) - K(P, Q)|^q dQ\right\}^{1/q}.$$

(The first of the above inequalities is a generalization of Minkowski's inequality [9, p. 148, §202].) Thus the conditions are sufficient to insure the complete continuity of $T\phi$. Now conversely if $T\phi$ is a completely continuous linear operation on L to L_q, it follows from 5.5 that the first and third condi-

tions are satisfied by some measurable kernel $K(P, Q)$. If $K(P, Q)$ is defined to be zero outside the square, then

$$\int_0^1 | K(P, Q+h) |^q dQ \leqq \int_0^1 | K(P, Q) |^q dQ,$$

and it follows again from 5.5 that

$$T_h\phi = \int_0^1 \{K(P, Q+h) - K(P, Q)\}\phi(P)dP$$

is a linear operation on L to L_q with

$$| T_h | = \text{ess. sup.} \left\{ \int_0^1 | K(P, Q+h) - K(P, Q) |^q dQ \right\}^{1/q}.$$

Since $T\phi$ is completely continuous, there is a $\delta_\epsilon > 0$ depending only upon an arbitrary positive number ϵ such that

$$\left\{ \int_0^1 | \psi(Q+h) - \psi(Q) |^q dQ \right\}^{1/q} < \epsilon \text{ for } | h | < \delta_\epsilon, \ \psi \text{ in } TS.$$

Fix h with $|h| < \delta_\epsilon$, then there is a ϕ in S such that

$$| T_h | \leqq \| T_h\phi \| + \epsilon;$$

that is,

$$\text{ess. sup.} \left\{ \int_0^1 | K(P, Q+h) - K(P, Q) |^q dQ \right\}^{1/q}$$

$$\leqq \left\{ \int_0^1 | \psi(Q+h) - \psi(Q) |^q dQ \right\}^{1/q} + \epsilon < 2\epsilon$$

which completes the proof of 6.5.

In view of the equivalence of $A.C.$ and L, 6.5 translates into

6.6. *The function $T\phi$ is a completely continuous linear operation on $A.C.$ to L_q $(1 < q < \infty)$ if and only if there is a measurable kernel $K(P, Q)$ which vanishes outside the unit square and satisfies the conditions*

$$\text{ess. sup.} \int_0^1 | K(P, Q) |^q dQ < \infty,$$

$$\lim_{h=0} \text{ ess. sup.} \int_0^1 | K(P, Q+h) - K(P, Q) |^q dQ = 0,$$

$$T\phi = \int_0^1 K(P, Q)\phi'(P)dP \text{ for } \phi \text{ in } A.C.$$

The norm of T is $|T| = \text{ess. sup.} \{\int_0^1 | K(P, Q) |^q dQ\}^{1/q}$.

For a linear operation on L to L_q $(1<q<\infty)$ the adjoint $\overline{T}$ on $\overline{L}_q$ to $\overline{L}$ may be regarded as a linear operation on $L_{q'}$ to M (since $\overline{L}_q$ and $\overline{L}$ are equivalent to $L_{q'}$ and M respectively). Recalling [1, p. 100, Theorems 3 and 4] that $|\overline{T}| = |T|$ and that $\overline{T}$ is completely continuous providing T is, the following theorem then follows from 5.5 and 6.5.

6.7. *Let $1<q<\infty$ and suppose the measurable kernel $K(P, Q)$ satisfies the condition*

$$\text{ess. sup.} \int_0^1 |K(P, Q)|^{q'} dQ < \infty,$$

then the operation

$$T\psi = \int_0^1 K(P, Q)\psi(Q)dQ$$

is a linear operation on L_q to M with $|T| = \text{ess. sup.} \{\int_0^1 |K(P, Q)|^{q'} dQ\}^{1/q'}$. The operation T is completely continuous provided

$$\lim_{h=0} \text{ess. sup.} \int_0^1 |K(P, Q+h) - K(P, Q)|^{q'} dQ = 0.$$

7. **Conclusion.** The problem of representation of operations on X to C (the space of continuous functions) where X is a Banach space for which the form of the general linear functional (numerically valued operator) is known, is much more easily handled than that of operations on L to L_q. This has been discussed in the case of C to C by J. Radon [11]. The fact that gives the hold on the problem is that for $\psi(Q)$ in C, $\psi(Q)=0$ for each Q provided $\|\psi\|=0$. This shows that by fixing Q the operator $\psi = T\phi$ on X to C can be interpreted as a functional. Thus, for example, in the case of L_p $(1<p<\infty)$ to C

$$\psi(Q) = \int_0^1 K(P, Q)\phi(P)dP, \tag{1}$$

where for each Q, $K(P, Q)$ is in $L_{p'}$ and thus

$$|T| \leqq \sup \left\{ \int_0^1 |K(P, Q)|^{p'} dP \right\}^{1/p'}.$$

But now for each ϕ in L_p, and each Q in $(0, 1)$

$$\lim_{Q' \to Q} \int_0^1 K(P, Q')\phi(P)dP = \int_0^1 K(P, Q)\phi(P)dP,$$

which is equivalent to saying that

$$\lim_{Q' \to Q} \int_0^t K(P, Q)dP = \int_0^t K(P, Q)dP, \qquad 0 \leqq t \leqq 1, \tag{2}$$

and

$$\sup \left\{ \int_0^1 | K(P, Q) |^{p'} dP \right\}^{1/p'} \equiv M < \infty. \tag{3}$$

Now fix Q so that

$$\left\{ \int_0^1 | K(P, Q) |^{p'} dP \right\}^{1/p'} \geqq M - \epsilon,$$

then there is a ϕ in L_p with $\|\phi\| = 1$ such that

$$| \psi(Q) | = \left| \int_0^1 K(P, Q)\phi(P)dP \right| \geqq M - \epsilon,$$

which shows that $\|\psi\| \geqq M$. Hence the norm of T is given by

$$| T | = \sup \left\{ \int_0^1 | K(P, Q) |^{p'} dP \right\}^{1/p'}. \tag{4}$$

To summarize: The function $\psi = T\phi$ is a linear operation on L_p $(1 < p < \infty)$ to C if and only if there exists a function $K(P, Q)$ in $L_{p'}$ for each Q and satisfying (1), (2), and (3). The norm of T is given by (4). Similar results can be obtained for other Banach spaces.

References

1. S. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932.
2. G. Birkhoff, *Integration of functions with values in a Banach space*, these Transactions, vol. 38 (1935), pp. 357–378.
3. S. Bochner, *Integration von Funktionen, deren Werte die Elemente eines Vektorraumes sind*, Fundamenta Mathematicae, vol. 20 (1933), pp. 262–276.
4. S. Bochner and J. von Neumann, *On compact solutions of operational differential equations*, I, Annals of Mathematics, (2), vol. 36 (1935), pp. 255–291.
5. J. A. Clarkson, *Uniformly convex spaces*, these Transactions, vol. 40 (1936), pp. 396–414.
6. N. Dunford, *Integration in general analysis*, these Transactions, vol 37 (1935), pp. 441–453.
7. N. Dunford and A. P. Morse, *Remarks on the preceding paper of James A. Clarkson*, these Transactions, vol. 40 (1936), pp. 415–420.
8. A. Haar, *Zur Theorie der Orthogonalen Funktionensysteme*, Inaugural-Dissertation, Göttingen, 1909, Chapter III, §§1–3.
9. Hardy, Littlewood and Pólya, *Inequalities*, Cambridge, 1934.
10. E. Hille and J. D. Tamarkin, *On the theory of linear integral equations*, II, Annals of Mathematics, (2), vol. 35 (1934), pp. 445–455.
11. J. Radon, *Über lineare Funktionaltransformationen und Funktionalgleichungen*, Sitzungsberichte, Akademie der Wissenschaften, Wien, Mathematische-Naturwissenschaftliche Klasse, vol. 128, 2a (1929), pp. 1083–1122.

12. M. Riesz, *Sur les ensembles compact de fonctions sommables*, Acta Szeged, vol. 6 (1933), pp. 136–142.

13. S. Saks, *Théorie de l'Intégrale*, Warsaw, 1933.

14. J. Schauder, *Eine Eigenschaft des Haarschen Orthogonalsystem*, Mathematische Zeitschrift, vol. 28 (1928), pp. 317–320.

15. W. Sierpiński, *Sur un problème concernant les ensembles mesurables superficiellement*, Fundamenta Mathematicae, vol. 1 (1920), pp. 112–115.

BROWN UNIVERSITY,
PROVIDENCE, R. I.

INITIAL MOTION IN FIELDS OF FORCE*

BY

AARON FIALKOW

1. **Introduction.** The motion of a particle in an arbitrary positional field of force suggests a number of problems in the differential geometry of plane curves. Several important families of curves are associated with every field of force. We here mention the following:

1. The totality of paths along which a particle starting from rest may move under the influence of the force. These ∞^2 curves are a subset of the ∞^3 *dynamical trajectories* of the field of force.

2. The family of curves whose direction at each point coincides with that of the force. These curves are the ∞^1 *lines of force.*

3. The system of curves on each of which the force has a constant direction. These are the ∞^1 *isoclinal curves* or *isoclines* of the field.

In this paper, we are concerned with certain geometric properties of these three families of curves. The theorems which are obtained are therefore a partial description of the photograph of the field of force.

Consider the motion of a particle which starts from rest at a point where the force is not zero. A first approximation to the path is given by the line of force on which the particle initially lies. However, the effect of inertia compels the particle to depart from this line of force and to travel in a straighter path. In general, the line of force and the trajectory will have the same initial direction but different initial curvatures. This situation and the problem of obtaining a suitable approximation to the beginning of the path were investigated by Kasner† and the writer‡ in previous papers. Kasner showed that *the curvature of the line of force is always three times the curvature of the trajectory*. When the initial curvature of the line of force is zero, Kasner studied the ratio between the infinitesimal departures of the line of force and of the path from their common tangent. We defined the set of these limiting ratios as the *ratio set*. Kasner obtained a theorem, assuming that the line of force has integral order of contact with its tangent, which we generalized to apply for any finite order of contact:

* Presented to the Society, April 10, 1936; received by the editors December 4, 1935.

† For a complete report of Kasner's work, see Proceedings of the National Academy of Sciences, vol. 20 (1934), pp. 130–136. Some results also appear in these Transactions, 1906–1910; Bulletin of the American Mathematical Society, vol. 16 (1909–1910), p. 172; *Differential Geometric Aspects of Dynamics*, American Mathematical Society Colloquium Publications, vol. 3, part 2, 1913, p. 9; Science, vol. 75 (1932), p. 671; Zurich Congress Proceedings, 1932, vol. 2, p. 180.

‡ These Transactions, vol. 38 (1935), pp. 89–105.

THEOREM A. *If the line of force has contact of order α with the tangent line, the trajectory will also have contact of order α; and the ratio set will be $2\alpha+1$.*

We also discussed the case of infinite order contact as well as some cases in which no definite order of contact existed and showed that the ratio set is always a closed positive interval which depends upon the contact between the line of force and its tangent. However, in the results that appear below, as well as in Theorem A, the ratio set always reduces to a single number; that is, there is a unique limiting ratio.

In the derivation of the formula for the ratio set, a special set of axes was employed which is also useful here. The initial point is taken as the origin of coordinates and the tangent to the line of force as the x-axis. An equation in x and y in which the variables are referred to the above set of axes has been called *normal*. The notion of the *direction function* at the initial point which we previously introduced also plays an important role in the present paper. If $D(x, y)$ is the direction of the force normalized with reference to a given initial point, then $y=D(x, 0)$ is the equation of the direction function associated with that point. The right-hand member represents the value of the slope at each point of the tangent line drawn to the line of force at the initial point. We proved a lemma concerning the line of force and the direction function to which we shall refer in subsequent proofs. This lemma may be stated as follows:

LEMMA. *The limit of the ratio of the infinitesimal departures of the line of force and of the curve*

$$y = \int_0^x D(x, 0)dx$$

from their common tangent is always unity.

It follows from this lemma and Theorem A that

THEOREM B. *The points at which the ratio set is $2n+1$ are identical with the points where the first $(n-1)$ derivatives* of the direction function vanish.*

The previous papers of Kasner and the writer investigated the geometric relationship between certain dynamical trajectories and lines of force. This relationship was stated in theorems about the ratio set. In this paper we describe the relationship between the ratio set and two new types of curves.

At certain points the ratio set is exceptional. These points usually form

* In our previous work, we assumed the existence of continuous first partial derivatives of the components of the field of force. At present we assume the existence and continuity of all derivatives which appear in the proofs.

a curve. We discuss some geometric properties of this curve and indicate how the ratio set depends upon the order of contact between this curve and its tangent.

The system of isoclinal curves of a field of force is geometrically determined by the lines of force. This leads to the conjecture that theorems similar to those previously derived for the lines of force apply in the case of the isoclines. The correctness of this guess is confirmed in our work. It is also shown that there is a close connection between the geometry of the system of iterated isoclinal curves and the ratio set.

2. **Ratio points and ratio curves.** It has been shown by Kasner that, in general, the ratio set is 3. In what follows, we discuss certain properties of the points at which the ratio set has an exceptional value. We call a point of a positional field at which the ratio set is not 3 a *ratio point*. A *ratio curve* is defined as a curve that consists entirely of ratio points. The fundamental relationship between the ratio points and the isoclines is given by

THEOREM I. *The ratio points are identical with the points at which any isoclinal curve of the field of force either is tangent to a line of force or has a singularity.*

Let $D(a, b)$ be the direction of the force at a ratio point (a, b). The components of the field are $\phi(x, y)$ and $\psi(x, y)$. In general, $\psi(a, b)/\phi(a, b)$ or $D(a, b)$ is not zero and a rotation and translation of axes is needed to write the equations in the normal form. In the following the subscript 1 is used to indicate normal coordinates and normal functions. The relations between the normal and given coordinates are

$$\begin{aligned} x &= x_1 \cos\theta - y_1 \sin\theta + a, \\ y &= x_1 \sin\theta + y_1 \cos\theta + b, \end{aligned} \tag{1}$$

where $\tan\theta = D(a, b)$. Similarly, we have

$$\begin{aligned} \phi_1(x_1, y_1) &= \phi(x, y)\cdot\cos\theta + \psi(x, y)\cdot\sin\theta, \\ \psi_1(x_1, y_1) &= -\phi(x, y)\cdot\sin\theta + \psi(x, y)\cdot\cos\theta, \end{aligned} \tag{2}$$

where x and y are to be replaced by their values given by (1). As stated in Theorem B, the ratio set at the initial point depends upon the behavior of the direction function $\psi_1(x_1, 0)/\phi_1(x_1, 0)$ or $D_1(x_1, 0)$. From (1) and (2), it follows that $D_1(x_1, 0)$ is equivalent to

$$\frac{\psi(x_1\cos\theta + a, x_1\sin\theta + b)\phi(a, b) - \phi(x_1\cos\theta + a, x_1\sin\theta + b)\psi(a, b)}{\psi(x_1\cos\theta + a, x_1\sin\theta + b)\psi(a, b) + \phi(x_1\cos\theta + a, x_1\sin\theta + b)\phi(a, b)}.$$

We assume without loss of generality that $\phi(a, b) \neq 0$ and divide both numera-

tor and denominator of the last expression by $\phi(x_1 \cos\theta+a, x_1 \sin\theta+b)\cdot\phi(a, b)$. Therefore

$$D_1(x_1, 0) = \frac{D(x_1 \cos\theta + a,\ x_1 \sin\theta + b) - D(a, b)}{D(x_1 \cos\theta + a,\ x_1 \sin\theta + b)D(a, b) + 1}. \tag{3}$$

Since (a, b) is a ratio point, by Theorem B, $(d/dx_1)D_1(x_1, 0)=0$. It follows from (3) that

$$D_x(a, b) + D(a, b)\cdot D_y(a, b) = 0. \tag{4}$$

Thus every ratio point is a solution of (4). The converse statement is also true.*

The equation of any isoclinal curve is $D(x, y)=c_1$. Differentiating, we obtain

$$D_x(x, y) + y'\cdot D_y(x, y) = 0 \tag{5}$$

which is the differential equation of the isoclinal curves. If an isocline is tangent to a line of force at (x_0, y_0), then $y'=D(x, y)$. Substituting in (5) and comparing with (4), we conclude that (x_0, y_0) is a ratio point. At a singularity of an isocline, $D_x=D_y=0$. From (4), it follows that such a point is also a ratio point. Since every ratio point lies on some isocline, the converse of the above statements is a consequence of the fact that $y'=D$ and $D_x=D_y=0$ are the only simultaneous solutions of (4) and (5). This concludes the proof.

The ratio curves are the solutions of (4). In general, a single ratio curve passes through a given ratio point. For this case, we have

THEOREM II. *Let a single ratio curve pass through a fixed ratio point in a direction which is not parallel to the force. Then the ratio set at this point is* 5. *Conversely, if the ratio set is* 5 *at a point through which there passes a unique ratio curve, this curve cuts the corresponding line of force.*

We suppose that the fixed ratio point is the origin and that the coordinates are normal. Differentiating (4), we obtain

$$D_{xx} + D_xD_y + DD_{xy} + (D_{xy} + D_y^2 + DD_{yy})\cdot y' = 0.$$

By the choice of coordinates $D(0, 0)$ is zero. Hence, by Theorem B, $D_x(0, 0)$ is also zero. Since there is a unique ratio curve through the origin, it follows that $D_{xy}+D_y^2+DD_{yy}\neq 0$ at the origin. Hence y' vanishes when and only when D_{xx} is zero. By Theorem B, we find that the ratio set is 5 only if $D_{xx}\neq 0$. Since by hypothesis, $y'\neq 0$ and the slope of the corresponding line of force vanishes, this result proves the theorem.

The analogue for ratio curves of Theorem A is

* If $D_x+DD_y\equiv 0$, all the lines of force are straight lines and every point is a ratio point.

THEOREM III. *Let a single ratio curve pass through a fixed ratio point in the direction of the force. Let the ratio curve have contact of order α with its tangent at this point. Then the line of force and the trajectory through the ratio point will both have contact of order $\alpha+2$ with their common tangent; and the ratio set will be $2\alpha+5$.*

We choose the coordinate axes as in the proof of the previous theorem. The implicit equation of the ratio curve is (4). By the theorem of the mean, this equation may be written as

$$[D_x(x, 0) + D(x, 0)\cdot D_y(x, 0)] + y\cdot[D_{xy}(x, \theta y) + D_y^2(x, \theta y) + D(x, \theta y)\cdot D_{yy}(x, \theta y)] = 0, \quad 0 < \theta < 1.$$

Representing the bracketed expressions by $E(x)$ and $F(x, y)$ respectively, the equation of the ratio curve is

$$E(x) + y\cdot F(x, y) = 0, \tag{6}$$

where by hypothesis, $F(0, 0)\neq 0$. Since the ratio curve has contact of order α,

$$\lim_{x\to 0} \frac{E(x)}{x^{\alpha+1}} = c_2 \qquad (\neq 0). \tag{7}$$

The line of force is a solution of $y'=D(x, y)$. Differentiating this equation, and applying the theorem of the mean, we obtain

$$y'' = E(x) + y\cdot F(x, y). \tag{8}$$

Let the line of force have generalized contact* of order β. We shall show that $\beta>\alpha$. Suppose $\beta\leqq\alpha$. Choose γ such that $\gamma+2>\beta>\gamma$. Dividing both members of (8) by $x^{\gamma+1}$ and letting x approach zero, we find that $\lim_{x\to 0} y''/x^{\gamma+1}=0$. Hence, by a repeated application of L'Hospital's Rule, the line of force has generalized contact of order $\geqq\gamma+2>\beta$. This contradiction shows that $\beta>\alpha$. We divide both members of (8) by $x^{\alpha+1}$ and let x approach zero. Since $\beta>\alpha$, by (7), $\lim_{x\to 0} y''/x^{\alpha+1}=c_2$. Therefore the line of force has ordinary contact of order $\alpha+2$; i.e., $\beta=\alpha+2$. The remainder of the theorem follows immediately from Theorem A.

* We defined generalized contact in our previous paper as follows:
A curve $y=f(x)$, where $f(x)$ is single-valued, continuous and

$$\lim_{x\to 0} \frac{f(x)}{x}=0$$

has generalized contact of order α with the x-axis if α is the upper bound of all numbers k such that

$$\lim_{x\to 0} \frac{f(x)}{x^{k+1}}=0.$$

α may be finite or infinite.

3. **Isoclinal curves.** The investigation of the isoclines of the field of force leads to results similar to those obtained for lines of force and ratio curves. From the point of view of the order of contact between one of these curves and its tangent at a fixed ratio point, the isoclines occupy an intermediate position between the lines of force and the ratio curves. Thus the analogue of Theorem A and Theorem III is

THEOREM IV. *Let a fixed point be a non-singular point of the isocline on which it lies. Let the isocline have contact of order α with the element of force at this point. Then the line of force and the trajectory through the point will both have contact of order $\alpha+1$ with their common tangent; and the ratio set will be $2\alpha+3$.*

The normal equation of the isoclinal curve is $D(x, y)=0$, or, using the theorem of the mean,

$$D(x, 0) + y \cdot D(x, \theta y) = 0, \qquad 0 < \theta < 1. \tag{9}$$

By hypothesis, $\lim_{x\to 0} y/x^{\alpha+1}=c_3$ $(\neq 0)$. Since the origin is not a singular point of the isocline, $D_y(0,0)\neq 0$. Therefore, by (9), $\lim_{x\to 0} D(x,0)/x^{\alpha+1}=c_4$ $(\neq 0)$. It follows from L'Hospital's Rule that the curve $y=\int_0^x D(x, 0)dx$ has contact of order $\alpha+1$ with the x-axis. Hence, by the lemma, the line of force must also have contact of order $\alpha+1$ with its tangent. The remainder of the theorem is a consequence of Theorem A.

A partially equivalent theorem for integral values of α exists. This theorem involves the following construction: The lines of force generate a family of isoclinal curves. These isoclinal curves in turn are the base of a second family of isoclines. By this process of iteration, any number of isoclinal families may be obtained if the components of the field of force have a sufficient number of continuous partial derivatives. For this configuration of iterated isoclinal families we have

THEOREM V. *Let the first n iterated isoclinal curves through a fixed point be tangent to the corresponding element of force. Let the $(n+1)$st iterated isoclinal curve pass through this point in a direction which is not parallel to the element of force. Then, if the fixed point is a non-singular point of these curves, the line of force and the trajectory through the point will both have contact of order $n+1$ with their common tangent; and the ratio set will be $2n+3$.*

This theorem is also an incomplete generalization of Theorem I. The proof follows. We assume the existence and continuity of all partial derivatives whose orders do not exceed $n+1$. The differential equation of the first family of isoclines is given by (5). Hence the second iterated isoclinal curves

are the solutions of $D_x+(y')_1 \cdot D_y=0$, where $(y')_1$ is any fixed value of y'. We differentiate this equation and obtain

$$D_{xx} + y' \cdot D_{xy} + (y')_1[D_{xy} + y'D_{yy}] = 0$$

as the differential equation of the family. By repeating this operation a sufficient number of times the differential equation of the kth iterated isoclinal curves can be obtained. It is easy to show that the form of this equation is

$$D_{x^{(k)}} + y' \cdot D_{x^{(k-1)}y} + G = 0. \tag{10}$$

G is a polynomial in y' and the partial derivatives of $D(x, y)$. Each term of G is multiplied by at least one $(y')_i$ $(i<k)$, where $(y')_i$ is a fixed value of the slope of the ith iterated isoclines. By hypothesis, since the coordinates are normal, the first n iterated isoclinal curves have zero slope at the origin. Hence zero must be substituted for $(y')_i$ $(i<n+1)$ in calculating the equations of the isoclines which pass through the origin. Letting $k=1, 2, \cdots, n$ in (10), we find that $D_{x^{(k)}}(0, 0)=0$. Similarly if $k=n+1$ in (10), G vanishes at the origin. By hypothesis, $(y')_{n+1}$ differs from zero at the origin. Therefore if $D_{x^{(n+1)}}(0, 0)=0$, the same is true for $D_{x^{(n)}y}(0, 0)$ and the origin is a singular point of the $(n+1)$st isocline. This contradiction shows that $D_{x^{(n+1)}}(0, 0)\neq 0$. It follows that the curve $y=\int_0^x D(x, 0)dx$ has contact of order $n+1$. The rest of the proof is similar to that of Theorem IV.

4. **Generalizations.** The results stated in Theorems III and IV may be extended to the case in which α denotes generalized instead of ordinary order of contact. Theorems similar to Theorem V of our previous paper are obtained. If the field of force is analytic, a result analogous to Theorem III exists when more than one ratio curve pass through the fixed ratio point. In this case, it follows from Weierstrass' Preparation Theorem for two variables that only a finite number of ratio curves can intersect at the point. With certain modifications the sum of their orders of contact replaces the single order of contact in Theorem III.

The curves upon which this study of the field of force is based are the lines of force, isoclines, and ratio curves. When may a curve be a member of more than one of these classes? The answer is simple. A curve cannot belong to two of these classes without being a member of the third. The only curves which are simultaneously lines of force, isoclines, and ratio curves are the straight lines.

COLUMBIA UNIVERSITY,
NEW YORK, N. Y.

TRANSACTIONS OF THE AMERICAN MATHEMATICAL SOCIETY

INDEX, VOLUMES 31–40
1929–1936

	Vol.	Pages
Adams, C. R. On the linear partial q-difference equation of general type	31_2	360–371
—— Note on integro-q-difference equations	31_4	861–867
—— On summability of double series	34_2	215–230
—— The space of functions of bounded variation and certain general spaces	40_3	421–438
Adams, C. R., and Clarkson, J. A. On definitions of bounded variation for functions of two variables	35_4	824–854
—— Properties of functions $f(x, y)$ of bounded variation	36_4	711–730
Agnew, R. P. The behavior of bounds and oscillations of sequences of functions under regular transformations	32_4	669–708
—— The effects of general regular transformations on oscillations of sequences of functions	33_2	411–424
—— On Riesz and Cesàro methods of summability	35_2	532–548
Albert, A. A. A determination of all normal division algebras in sixteen units	31_2	253–260
—— New results in the theory of normal division algebras	32_2	171–195
—— On direct products, cyclic division algebras, and pure Riemann matrices	33_1 33_4	219–234 999
—— On normal division algebras of type R in thirty-six units	33_1	235–243
—— On direct products	33_3	690–711
—— Normal division algebras of degree four over an algebraic field	34_2	363–372
—— On normal simple algebras	34_3	620–625
—— Non-cyclic algebras of degree and exponent four	35_1	112–121
—— Cyclic fields of degree eight	35_4	949–964
—— Normal division algebras over a modular field	36_2	388–394
—— On normal Kummer fields over a non-modular field	36_4	885–892
—— On cyclic fields	37_3	454–462
—— Normal division algebras of degree p^e over F of characteristic p	39_1	183–188
—— Simple algebras of degree p^e over a centrum of characteristic p	40_1	112–126
Albert, A. A., and Hasse, H. A determination of all normal division algebras over an algebraic number field	34_3	722–726

	Vol.	Pages
Ayres, W. L. Concerning the arc-curves and basic sets of a continuous curve, second paper	31_3	595–612
—— On the regular points of a continuum	33_1	252–262
Bamforth, F. R. Surface transformations	32_3	417–450
Bamforth, F. R., and Birkhoff, G. D. Divergent series and singular points of ordinary differential equations	32_1	114–146
Bartky, W., and MacMillan, W. D. Permanent configurations in the problem of four bodies	34_4	838–875
Basye, R. E. Simply connected sets	38_2	341–356
Bateman, H. The *k*-function, a particular case of the confluent hypergeometric function	33_4	817–831
Beckenbach, E. F., and Radó, T. Subharmonic functions and minimal surfaces	35_3	648–661
—— Subharmonic functions and surfaces of negative curvature	35_3	662–674
Bell, E. T. Certain invariant sequences of polynomials	31_3	405–421
—— Numbers of representations of integers in a certain triad of ternary quadratic forms	32_1	1– 5
—— Analytic functions in the irregular field of all numerical functions	32_4	869–878
—— Singular relations between certain arithmetical functions	33_1	65– 71
—— Arithmetical composition and inversion of functions over classes	33_4	897–933
—— The Latin square, or cyclic, functions	35_3	734–745
—— Polynomial diophantine systems	35_4	903–914
—— On the power series for elliptic functions	36_4	841–852
—— General relations between Bernoulli, Euler, and allied polynomials	38_3	493–500
—— A functional equation in arithmetic	39_3	341–344
Bernstein, B. A. A set of four postulates for Boolean algebra in terms of the "implicative" operation	36_4	876–884
Binney, J. H. An elliptic system of integral equations on summable functions	37_2	254–265
Bird, M. T., Carmichael, R. D., and Martin, W. T. On a classification of integral functions by means of certain invariant point properties	40_3	462–473
Birkhoff, G. Integration of functions with values in a Banach space	38_2	357–378
Birkhoff, G., and Hall, P. On the order of groups of automorphisms	39_3	496–499
Birkhoff, G. D., and Bamforth, F. R. Divergent series and singular points of ordinary differential equations	32_1	114–146

	Vol.	Pages
Black, A. H. Types of involutorial space transformations associated with certain rational curves	34_4	795–810
Bliss, G. A., and Hestenes, M. R. Sufficient conditions for a problem of Mayer in the calculus of variations	35_1	305–326
Boas, R. P. Some theorems on Fourier transforms and conjugate trigonometric integrals	40_2	287–308
Bochner, S. Summation of multiple Fourier series by spherical means	40_2	175–207
Bochner, S., and von Neumann, J. Almost periodic functions in groups, II	37_1	21– 50
Bosanquet, L. S. Some arithmetic means connected with Fourier series	39_2	189–204
Brahana, H. R. Groups $\{S, T\}$ whose commutator subgroups are abelian	35_2	386–396
—— Metabelian groups of order p^{n+m} with commutator subgroups of order p^m	36_4	776–792
—— Note on irreducible quartic congruences	38_2	395–400
Bristow, L. Expansion theory associated with linear differential equations and their regular singular points	33_2	455–474
Brown, A. B. Functional dependence	38_2	379–394
Brown, A. B., and Koopman, B. O. On the covering of analytic loci by complexes	34_2	231–251
—— The Riemann multiple-space and algebroid functions	36_3	618–626
Bullock, R. C. Non-conjugate osculating quadrics of a curve on a surface	35_2	518–531
Burington, R. S. On the equivalence of quadrics in m-affine n-space and its relation to the equivalence of $2m$-pole networks	38_1	163–176
Cairns, S. S. An axiomatic basis for plane geometry	35_1	234–244
—— The generalized theorem of Stokes	40_1	167–174
Calugaréano, G. Les fonctions polygènes comme intégrales d'équations différentielles	31_2	372–378
—— On differential equations admitting polygenic integrals	32_1	110–113
Cameron, R. H. Almost periodic transformations	36_2	276–291
Camp, B. H. The multinomial solid and the chi test	31_1	133–144
Carlitz, L. On Galois fields of certain types	32_3	451–472
—— On abelian fields	35_1	122–136
—— On the representation of a polynomial in a Galois field as the sum of an even number of squares	35_2	397–410
Carmichael, R. D. Systems of linear difference equations and expansions in series of exponential functions	35_1	1– 28
Carmichael, R. D., Martin, W. T., and Bird, M. T. On a		

	Vol.	Pages
classification of integral functions by means of certain invariant point properties	40_3	462–473
Carpenter, A. F. Tetrads of ruled surfaces	31_3	523–528
Carson, J. R. The asymptotic solution of an operational equation	31_4	782–792
Chittenden, E. W. On general topology and the relation of the properties of the class of all continuous functions to the properties of space	31_2	290–321
Chokhate, J. See Shohat, J. A.		
Church, A., and Rosser, J. B. Some properties of conversion	39_3	472–482
Churchill, R. V. On the geometry of the Riemann tensor	34_1	126–152
—— Canonical forms for symmetric linear vector functions in pseudo-euclidean space	34_4	784–794
Clarkson, J. A. Uniformly convex spaces	40_3	396–414
Clarkson, J. A., and Adams, C. R. On definitions of bounded variation for functions of two variables	35_4	824–854
—— Properties of functions $f(x, y)$ of bounded variation	36_4	711–730
Cleveland, C. M. On the existence of acyclic curves satisfying certain conditions with respect to a given continuous curve	33_4	958–978
Coe, C. J. Exterior motion in the restricted problem of three bodies	34_4	811–837
Cook, A. J. Pairs of rectilinear congruences with generators in one-to-one correspondence	32_1	31– 46
Craig, H. V. On parallel displacement in a non-Finsler space	33_1	125–142
Cramlet, C. M. A complete system of tensors of linear homogeneous second-order differential equations	34_3	626–644
Currier, A. E. The variable end point problem of the calculus of variations including a generalization of the classical Jacobi condition	34_3	689–704
—— Proof of the fundamental theorems on second-order cross partial derivatives	35_1	245–253
Curtiss, J. Interpolation in regularly distributed points	38_3	458–473
Cutler, E. H. On the curvatures of a curve in Riemann space	33_4	832–838
—— Frenet formulas for a general subspace of a Riemann space	33_4	839–850
Davis, D. R. Integrals whose extremals are a given $2n$-parameter family of curves	33_1	244–251
Diamond, A. H. The complete existential theory of the Whitehead-Huntington set of postulates for the algebra of logic	35_4	940–948
—— Correction to a paper on the Whitehead-Huntington postulates	36_4	893
Dickson, L. E. Universal quadratic forms	31_1	164–189
—— Constructions of division algebras	32_2	319–334

	Vol.	Pages
—— Waring's problem for cubic functions	36_1	1– 12
—— A new method for Waring theorems with polynomial summands	36_4	731–748
—— Cyclotomy and trinomial congruences	37_3	363–380
—— Cyclotomy when e is composite	38_2	187–200
—— A new method for Waring theorems with polynomial summands, II	39_2	205–208
Doob, J. L. The boundary values of analytic functions	34_1	153–170
—— The boundary values of analytic functions. II	35_2	418–451
—— Probability and statistics	36_4	759–775
—— Statistical estimation	39_3	410–421
Douglas, J. Solution of the problem of Plateau	33_1	263–321
—— One-sided minimal surfaces with a given boundary	34_4	731–756
Dunford, N. Integration in general analysis	37_3	441–453
—— Corrections to the paper "Integration in general analysis"	38_3	600–601
—— Integration and linear operations	40_3	474–494
Dunford, N., and Morse, A. P. Remarks on the preceding paper of James A. Clarkson	40_3	415–420
Dye, L. A. Involutorial transformations in S_3 of order n with an $(n-1)$-fold line	32_2	251–261
Dye, L. A., and Sharpe, F. R. The Bertini transformation in space	36_2	292–305
Earl, J. M. Polynomials of best approximation on an infinite interval	32_4	888–904
Eiesland, J. The ruled V_4^4 in S_5 associated with a Schläfli hexad	36_2	315–326
Emch, A. Triple and multiple systems, their geometric configurations and groups	31_1	25– 42
Engstrom, H. T. On the common index divisors of an algebraic field	32_2	223–237
—— The theorem of Dedekind in the ideal theory of Zolotarev	32_4	879–887
—— On sequences defined by linear recurrence relations	33_1	210–218
Evans, G. C. On potentials of positive mass. Part I	37_2	226–253
—— Potentials of positive mass. II	38_2	201–236
Fejér, L. Trigonometrische Reihen und Potenzreihen mit mehrfach monotoner Koeffizientenfolge	39_1	18– 59
Fekete, M. Some generalizations of Paley's theorems on Fourier series with positive coefficients	38_2	237–249
Feraud, L. On Birkhoff's Pfaffian systems	32_4	817–831
Fialkow, A. Trajectories and lines of force	38_1	89–105
—— Initial motion in fields of force	40_3	495–501
Ford, L. R. The limit points of a group	31_4	821–828
Fort, T. Generalized factorial series	31_2	233–240

	Vol.	Pages
Fréchet, M., and Shohat, J. A proof of the generalized second limit-theorem in the theory of probability	33_2 / 33_4	533–543 / 999
Gehman, H. M. On extending a continuous (1-1) correspondence (second paper)	31_2	241–252
Gergen, J. J. Convergence criteria for double Fourier series	35_1	29– 63
Gergen, J. J., and Littauer, S. B. Continuity and summability for double Fourier series	38_3	401–435
Geronimus, J. On orthogonal polynomials	33_1	322–328
Ghosh, S. On the solution of the equations of elastic equilibrium suitable for elliptic boundaries	32_1	47– 60
Gillespie, D. C., and Hurwitz, W. A. On sequences of continuous functions having continuous limits	32_3	527–543
Gore, G. D. Inscribed sequences of surfaces associated with generalized sequences of Laplace	36_3	530–541
Gourin, E. On irreducible polynomials in several variables which become reducible when the variables are replaced by powers of themselves	32_3	485–501
Graustein, W. C. Parallelism and equidistance in classical differential geometry	34_3	557–593
—— The geometry of Riemannian spaces	36_3	542–585
Graves, L. M. A transformation of the problem of Lagrange in the calculus of variations	35_3	675–682
—— The existence of an extremum in problems of Mayer	39_3	456–471
Griffin, M. Invariants of Pfaffian systems	35_4	929–939
Grove, V. G. Nets with equal W invariants	31_4	845–852
—— Contributions to the general theory of transformations of nets	31_4	853–860
—— A canonical form of Green's projective analogue of the Gauss differential equations	32_3	473–484
—— The transformation E of nets	33_1	147–152
—— The transformation C of nets in hyperspace	33_3	733–741
—— Contributions to the theory of transformations of nets in a space S_n	35_3	683–688
—— On a certain correspondence between surfaces in hyperspace	36_3	627–636
—— Differential geometry of a certain type of surface in S_4	39_1	60– 70
—— Differential geometry of a surface at a planar point	40_1	155–166
Hall, P., and Birkhoff, G. On the order of groups of automorphisms	39_3	496–499
Harshbarger, F. The geometric configuration defined by a special algebraic relation of genus four	33_2	557–578

	Vol.	Pages
Hasse, H. Theory of cyclic algebras over an algebraic number field	34_1	171–214
—— Additional note to the author's "Theory of cyclic algebras over an algebraic number field"	34_3	727–730
Hasse, H., and Albert, A. A. A determination of all normal division algebras over an algebraic number field	34_3	722–726
Hazlett, O. C. Homogeneous polynomials with a multiplication theorem	31_2	223–232
—— On division algebras	32_4	912–925
Hedlund, G. A. Poincaré's rotation number and Morse's type number	34_1	75– 97
Heineman, E. R. Generalized Vandermonde determinants	31_3	464–476
Herzog, F. Systems of algebraic mixed difference equations	37_2	286–300
Hestenes, M. R. Sufficient conditions for the general problem of Mayer with variable end points	35_2	479–490
—— Sufficient conditions for the problem of Bolza in the calculus of variations	36_4	793–818
Hestenes, M. R., and Bliss, G. A. Sufficient conditions for a problem of Mayer in the calculus of variations	35_1	305–326
Hickson, A. O. An application of the calculus of variations to boundary value problems	31_3	563–579
Hildebrandt, T. H. On the interchange of limit and Lebesgue integral for a sequence of functions	33_2	441–443
—— On bounded linear functional operations	36_4	868–875
Hille, E. Notes on linear transformations, I	39_1	131–153
Hille, E., and Tamarkin, J. D. On the summability of Fourier series. I	34_4	757–783
Hinrichsen, J. J. L. On the problem of n bodies	36_2	306–314
Hollcroft, T. R. The general web of algebraic surfaces of order n and the involution defined by it	35_4	855–868
Hopf, E. Theory of measure and invariant integrals	34_2	373–393
—— Fuchsian groups and ergodic theory	39_2	299–314
Hopkins, C. Metabelian groups of order p^m, $p>2$	37_1	161–195
Hotelling, H. The consistency and ultimate distribution of optimum statistics	32_4	847–859
Hull, R. The numbers of solutions of congruences involving only kth powers	34_4	908–937
—— Maximal orders in rational cyclic algebras of odd prime degree	38_3	515–530
—— The maximal orders of generalized quaternion division algebras	40_1	1– 11

	Vol.	Pages
Huntington, E. V. New sets of independent postulates for the algebra of logic, with special reference to Whitehead and Russell's *Principia Mathematica*	35_1	274–304
—— Boolean algebra. A correction	35_2	557–558
—— A second correction	35_4	971
—— Inter-relations among the four principal types of order	38_1	1– 9
Hurwitz, W. A., and Gillespie, D. C. On sequences of continuous functions having continuous limits	32_3	527–543
Hutchinson, J. I. Properties of functions represented by the Dirichlet series $\sum(av+b)^{-s}$, or by linear combinations of such series	31_2	322–344
Ingraham, M. H., and Wegner, K. W. The equivalence of pairs of Hermitian matrices	38_1	145–162
Jackson, D. A note on closest approximation	31_2	215–222
—— Bernstein's theorem and trigonometric approximation	40_2	225–251
James, R. D. The value of the number $g(k)$ in Waring's problem	36_2	395–444
Jeffery, R. L. The integrability of a sequence of functions	33_2	433–440
—— Non-absolutely convergent integrals with respect to functions of bounded variation	34_3	645–675
—— Sets of k-extent in n-dimensional space	35_3	629–647
—— Derived numbers with respect to functions of bounded variation	36_4	749–758
Jessen, B., and Wintner, A. Distribution functions and the Riemann zeta function	38_1	48– 88
Jones, B. W. A new definition of genus for ternary quadratic forms	33_1	92–110
—— The regularity of a genus of positive ternary quadratic forms	33_1	111–124
Kellogg, O. D. On the derivatives of harmonic functions on the boundary	33_2	486–510
—— Converses of Gauss' theorem on the arithmetic mean	36_2	227–242
Ketchum, G. S. On certain generalizations of the Cauchy-Taylor expansion theory	40_2	208–224
Kogbetliantz, E. Contribution à l'étude du saut d'une fonction donnée par son developpement en série d'Hermite ou de Laguerre	38_1	10– 47
Koopman, B. O. On the singularities of linear partial differential equations	31_4	687–693
—— On distributions admitting a sufficient statistic	39_3	399–409
Koopman, B. O., and Brown, A. B. On the covering of analytic loci by complexes	34_2	231–251
—— The Riemann multiple-space and algebroid functions	36_3	618–626

	Vol.	Pages
Lamson, K. W. Some differential and algebraic consequences of the Einstein field equations	32_4	709–722
Lane, E. P. Hypergeodesic mapping of a surface on a plane	32_3	558–568
—— Integral surfaces of pairs of partial differential equations of the third order	32_4	782–793
—— Surfaces and curvilinear congruences	34_3	676–688
—— The moving trihedron	36_3	696–710
—— A canonical power series expansion for a surface	37_3	463–482
Langer, R. E. The boundary problem associated with a differential equation in which the coefficient of the parameter changes sign	31_1	1– 24
—— The asymptotic location of the roots of a certain transcendental equation	31_4	837–844
—— The expansion problem in the theory of ordinary linear differential systems of the second order	31_4	868–906
—— The boundary problem associated with a differential system rational in the parameter	32_2	238–250
—— On the asymptotic solutions of ordinary differential equations, with an application to the Bessel functions of large order	33_1	23– 64
—— On the asymptotic solutions of differential equations, with an application to the Bessel functions of large complex order	34_3	447–480
—— The asymptotic solutions of certain linear ordinary differential equations of the second order	36_1	90–106
—— The solutions of the Mathieu equation with a complex variable and at least one parameter large	36_3	637–695
—— On the asymptotic solutions of ordinary differential equations, with reference to the Stokes' phenomenon about a singular point	37_3	397–416
LaPaz, L. An inverse problem of the calculus of variations for multiple integrals	32_3	509–519
Laporte, O., and Rainich, G. Y. Stereographic parameters and pseudo-minimal hypersurfaces	39_1	154–182
Latimer, C. G. On certain identities in theta functions	32_4	832–846
—— On the class number of a cyclic field	35_2	411–417
—— On ideals in generalized quaternion algebras and Hermitian forms	38_3	436–446
—— On the class number of a quaternion algebra with a negative fundamental number	40_2	318–323
—— On ideals in a quaternion algebra and the representation of integers by Hermitian forms	40_3	439–449

	Vol.	Pages
Lefschetz, S., and Whitehead, J. H. C. On analytical complexes	35_2	510–517
Lehmer, D. H. Arithmetic of double series	33_4	945–957
Lehmer, D. N. On the congruences connected with certain magic squares	31_3	529–551
Leighton, W., and Morse, M. Singular quadratic functionals	40_2	252–286
Lev, J. Effects of linear transformations on the divergence of bounded sequences and functions	35_4	888–896
Lewis, D. C. Infinite systems of ordinary differential equations with applications to certain second-order partial differential equations	35_4	792–823
Lewy, H. A priori limitations for solutions of Monge-Ampère equations	37_3	417–434
Littauer, S. B., and Gergen, J. J. Continuity and summability for double Fourier series	38_3	401–435
Litzinger, M. A basis for residual polynomials in n variables	37_2	216–225
Lorch, E. R., and Riesz, F. The integral representation of unbounded self-adjoint transformations in Hilbert space	39_2	331–340
Lubben, R. G. Separation theorems with applications to questions concerning accessibility and plane continua	31_3	503–522
McCoy, N. H. On commutation formulas in the algebra of quantum mechanics	31_4	793–806
—— On the resultant of a system of forms homogeneous in each of several sets of variables	35_1	215–233
—— On quasi-commutative matrices	36_2	327–340
—— Quasi-commutative rings and differential ideals	39_1	101–116
McEwen, W. H. Problems of closest approximation connected with the solution of linear differential equations	33_4	979–997
McKinsey, J. C. C. On Boolean functions of many variables	40_3	343–362
McShane, E. J. Parametrizations of saddle surfaces, with application to the problem of Plateau	35_3	716–733
—— Existence theorems for double integral problems of the calculus of variations	38_3	549–563
MacColl, L. A. On the distributions of the zeros of sums of exponentials of polynomials	36_2	341–360
MacDuffee, C. C. An introduction to the theory of ideals in linear associative algebras	31_1	71– 90
—— The discriminant matrix of a semi-simple algebra	33_2	425–432
—— Covariants of r-parameter groups	39_1	71– 82
MacGregor, C. W. The potential function method for the solution of two-dimensional stress problems	38_1	177–186

	Vol.	Pages
MacLane, S. A construction for absolute values in polynomial rings	40_3	363–395
MacMillan, W. D., and Bartky, W. Permanent configurations in the problem of four bodies	34_4	838–875
MacQueen, M. L. A projective generalization of metrically defined associate surfaces	36_4	826–840
Maier, W. E. Elementary properties of the l_ν-functions	32_4 33_4	905–911 998
Manning, D. On simply transitive groups with transitive Abelian subgroups of the same degree	40_2	324–342
Manning, W. A. The degree and class of multiply transitive groups, II	31_4	643–653
—— The degree and class of multiply transitive groups, III	35_3	585–599
Marden, M. On the zeros of linear partial fractions	32_1	81–109
—— On the zeros of certain rational functions	32_4	658–668
—— On Stieltjes polynomials	33_4	934–944
Martin, W. T., Bird, M. T., and Carmichael, R. D. On a classification of integral functions by means of certain invariant point properties	40_3	462–473
Mayer, W. Die differential Geometrie der Untermannigfaltigkeiten des R_n konstanter Krümmung	38_2	267–309
Menger, K. Remarks concerning the paper of W. L. Ayres on the regular points of a continuum	33_3	663–667
Mickelson, E. L. On the approximate representation of a function of two variables	33_3	759–781
Miles, E. R. C. Boundary value problems for potentials of a single layer	31_1	190–203
Miller, G. A. Groups in which every operator has at most a prime number of conjugates	35_4	897–902
—— Groups in which the squares of the elements are a dihedral subgroup	36_4	819–825
Milne, W. E. On the degree of convergence of the Gram-Charlier series	31_3	422–443
—— On the degree of convergence of expansions in an infinite interval	31_4	907–918
—— On the maximum absolute value of the derivative of $e^{-x^2}P_n(x)$	33_1	143–146
Montgomery, D. Sections of point sets	35_4	915–928
Montgomery, D., and Zippin, L. Periodic one-parameter groups in three-space	40_1	24– 36
Moore, C. N. On Gibbs's phenomenon for the developments in Bessel's functions	32_3	409–416

	Vol.	Pages
Morse, A. P., and Dunford, N. Remarks on the preceding paper of James A. Clarkson	40_3	415–420
Morse, M. The foundations of the calculus of variations in the large m-space (first paper)	31_3	379–404
—— The foundations of a theory of the calculus of variations in the large in m-space (second paper)	32_4	599–631
—— The critical points of a function of n variables	33_1	72– 91
—— Sufficient conditions in the problem of Lagrange without assumptions of normalcy	37_1	147–160
Morse, M., and Leighton, W. Singular quadratic functionals	40_2	252–286
Murray, F. J. Linear transformations between Hilbert spaces and the application of this theory to linear partial differential equations	37_2	301–338
—— Linear transformations in $\mathfrak{L}_p$, $p>1$	39_1	83–100
Myers, S. B. Sufficient conditions in the problem of the calculus of variations in n-space in parametric form and under general end conditions	35_3	746–760
von Neumann, J. Almost periodic functions in a group. I	36_3	445–492
—— On complete topological spaces	37_1	1– 20
von Neumann, J., and Bochner, S. Almost periodic functions in groups, II	37_1	21– 50
Newton, A. V. Consecutive covariant configurations at a point of a space curve	36_1	44– 62
Offord, A. C. On Fourier transforms. III	38_2	250–266
Oldenburger, R. Non-singular multilinear forms and certain p-way matrix factorizations	39_3	422–455
Ore, O. On a special class of polynomials	35_3	559–584
—— Contributions to the theory of finite fields	36_2	243–274
—— Errata in my paper "On a special class of polynomials"	36_2	275
Paley, R. E. A. C. A special integral function	35_3	709–715
Paley, R. E. A. C., and Wiener, N. Notes on the theory and application of Fourier transforms. I–II	35_2	348–355
—— Notes on the theory and application of Fourier transforms. III, IV, V, VI, VII	35_4	761–791
Pall, G. On sums of two or four values of a quadratic function of x	34_1	98–125
—— The structure of the number of representations function in a binary quadratic form	35_2	491–509
Perkins, F. W. The Dirichlet problem for domains with multiple boundary points	38_1	106–144
Pierpont, J. Foucault's pendulum in elliptic space	31_3	444–447
Poor, V. C. Residues of polygenic functions	32_2	216–222

	Vol.	Pages
Poritsky, H. On certain polynomial and other approximations to analytic functions	34_2	274–331
Post, E. L. Generalized differentiation	32_4	723–781
Prenowitz, W. The characterization of plane collineations in terms of homologous families of lines	38_3	564–599
Price, G. B. On reversible dynamical systems	37_1	51– 79
Rademacher, H. On prime numbers of real quadratic fields in rectangles	39_3	380–398
Radó, T. An iterative process in the problem of Plateau	35_4	869–887
—— On convex functions	37_2	266–285
Radó, T., and Beckenbach, E. F. Subharmonic functions and minimal surfaces	35_3	648–661
—— Subharmonic functions and surfaces of negative curvature	35_3	662–674
Rainich, G. Y., and Laporte, O. Stereographic parameters and pseudo-minimal hypersurfaces	39_1	154–182
Randolph, J. F. Carathéodory measure and a generalization of the Gauss-Green lemma	38_3	531–548
Ranum, A. The singular points of analytic space-curves	31_1 31_4	145–163 931
Raudenbush, H. W. Ideal theory and algebraic differential equations	36_2	361–368
Regan, F. The application of the theory of admissible numbers to time series with constant probability	36_3	511–529
Reid, W. T. Properties of solutions of an infinite system of ordinary linear differential equations of the first order with auxiliary boundary conditions	32_2	284–318
—— Expansion problems associated with a system of integral equations	33_2	475–485
Riesz, F., and Lorch, E. R. The integral representation of unbounded self-adjoint transformations in Hilbert space	39_2	331–340
Ritt, J. F. Algebraic combinations of exponentials	31_4 32_4	654–679 944
—— On the zeros of exponential polynomials	31_4	680–686
—— On a certain ring of functions of two variables	32_2	155–170
—— Manifolds of functions defined by [illegible] algebraic differential equations	32_4	569–598
Roberts, J. H. Concerning non-dense plane continua	32_1	6– 30
—— Concerning topological transformations in E_n	34_2	252–262
Roos, C. F. A general problem of minimizing an integral with discontinuous integrand	31_1	58– 70
Rosser, J. B., and Church, A. Some properties of conversion	39_3	472–482

	Vol.	Pages
Rosskopf, M. F. Some inequalities for non-uniformly bounded ortho-normal polynomials	36_4	853–867
Roth, W. E. On the unilateral equation in matrices	32_1	61– 80
—— On the equation $P(A, X)=0$ in matrices	35_3	689–708
—— On the characteristic values of the matrix $f(A, B)$	39_2	234–243
—— On k-commutative matrices	39_3	483–495
Rowe, C. H. A kinematical treatment of some theorems on normal rectilinear congruences	31_4	919–930
Rupp, C. A. An extension of Pascal's theorem	31_3	580–594
Russell, H. G., and Walsh, J. L. On the convergence and overconvergence of sequences of polynomials of best simultaneous approximation to several functions analytic in distinct regions	36_1	13– 28
Rutledge, G. The basic power series of interpolation	31_4	807–820
Rutt, N. E. On certain types of plane continua	33_3	806–816
Saks, S. On some functionals	35_2	549–556
—— Addition to the note on some functionals	35_4	965–970
Scheffé, H. Asymptotic solutions of certain linear differential equations in which the coefficient of the parameter may have a zero	40_1	127–154
Scherberg, M. G. The degree of convergence of a series of Bessel functions	35_1	172–183
Schoenberg, I. J. On finite-rowed systems of linear inequalities in infinitely many variables	34_3	594–619
—— On finite-rowed systems of linear inequalities in infinitely many variables. II	35_2	452–478
—— On asymptotic distributions of arithmetical functions	39_2	315–330
—— On the zeros of successive derivatives of integral functions	40_1	12– 23
Schwid, N. The asymptotic forms of the Hermite and Weber functions	37_2	339–362
Seidel, W. On the cluster values of analytic functions	34_1	1– 21
—— On the distribution of values of bounded analytic functions	36_1	201–226
Sharpe, F. R., and Dye, L. A. The Bertini transformation in space	36_2	292–305
Sheffer, I. M. Expansions in generalized Appell polynomials, and a class of related linear functional equations	31_2	261–280
—— Systems of infinitely many linear differential equations of infinite order, with constant coefficients	31_2	281–289
—— On the properties of polynomials satisfying a linear differential equation: Part I	35_1	184–214
—— A local solution of the difference equation $\Delta y(x)=F(x)$ and of related equations	39_3	345–379

	Vol.	Pages
Sherman, J. On the numerators of the convergents of the Stieltjes continued fractions	35_1	64– 87
Shohat, J. A. (Chokhate, J.). On a certain formula of mechanical quadratures with non-equidistant ordinates	31_3	448–463
Shohat, J., and Fréchet, M. A proof of the generalized second limit-theorem in the theory of probability	33_2 33_4	533–543 999
Sibert, H. W. Moderately thick circular plates with plane faces	33_2	329–369
Siegel, C. L. The volume of the fundamental domain for some infinite groups	39_2	209–218
—— Mittelwerte arithmetischer Funktionen in Zahlkörpern	39_2	219–224
—— Über die algebraischen Integrale des restringierten Dreikörperproblems	39_2	225–233
Simmons, H. A. Classes of maximum numbers and minimum numbers that are associated with certain symmetric equations in n reciprocals	34_4	876–907
—— The first and second variations of an n-tuple integral in the case of variable limits	36_1	29– 43
Singer, J. Three-dimensional manifolds and their Heegaard diagrams	35_1	88–111
Sinkov, A. Families of groups generated by two operators of the same order	35_2	372–385
Snyder, V. On a series of involutorial Cremona transformations of space defined by a pencil of ruled surfaces	35_2	341–347
Sokolnikoff, I. S. On a solution of Laplace's equation with an application to the torsion problem for a polygon with reentrant angles	33_3	719–732
Stephens, R. Continuous transformations of abstract spaces	34_2	394–408
Stokes, R. W. A geometric theory of solution of linear inequalities	33_3	782–805
Stone, M. H. The theory of representations for Boolean algebras	40_1	37–111
Sullivan, M. M. On the derivatives of Newtonian and logarithmic potentials near the acting masses	35_1	137–171
Suschkewitsch, A. On a generalization of the associative law	31_1	204–214
Swingle, P. M. A certain type of continuous curve and related point sets	33_2	544–556
Synge, J. L. The apsides of general dynamical systems	34_3	481–522
Szász, O. Convergence properties of Fourier series	37_3	483–500
—— Converse theorems of summability for Dirichlet's series	39_1	117–130
Szegö, G. A problem concerning orthogonal polynomials	37_1	196–206
—— Inequalities for the zeros of Legendre polynomials and related functions	39_1	1– 17

	Vol.	Pages
—— Correction to the paper "A problem concerning orthogonal polynomials"	39_3	500
—— On some Hermitian forms associated with two given curves of the complex plane	40_3	450–461
Tamarkin, J. D. The notion of the Green's function in the theory of integro-differential equations. II	32_4	860–868
—— On a theorem of S. Bernstein-Widder	33_4	893–896
Tamarkin, J. D., and Hille, E. On the summability of Fourier series. I	34_4	757–783
Thomas, J. M. Matrices of integers ordering derivatives	33_2	389–410
—— The condition for an orthonomic differential system	34_2	332–338
—— Pfaffian systems of species one	35_2	356–371
Thomas, T. Y. Algebraic characterizations in complex differential geometry	38_3	501–514
Trjitzinsky, W. J. On composition of singularities	32_2	196–215
—— Laplace integrals and factorial series in the theory of linear differential and linear difference equations	37_1	80–146
Underwood, R. S. On universal quadratic null forms in five variables	33_3	742–758
Uspensky, J. V. A method for finding units in cubic orders of a negative discriminant	33_1	1– 22
—— On the expansion of the remainder in the Newton-Cotes formula	37_3	381–396
Vaidyanathaswamy, R. The theory of multiplicative arithmetic functions	33_2	579–662
Vandiver, H. S. On Fermat's last theorem	31_4 33_4	613–642 998
—— On power characters of singular integers in a properly irregular cyclotomic field	32_3	391–408
Wall, H. S. On the Padé approximants associated with the continued fraction and series of Stieltjes	31_1	91–116
—— On extended Stieltjes series	31_4	771–781
—— On the Padé approximants associated with a positive definite power series	33_2	511–532
—— General theorems on the convergence of sequences of Padé approximants	34_2	409–416
Walsh, J. L. Note on the expansion of analytic functions in series of polynomials and in series of other analytic functions	31_1	53– 57
—— On approximation by rational functions to an arbitrary function of a complex variable	31_3	477–502

	Vol.	Pages
—— Boundary values of an analytic function and the Tchebycheff method of approximation	32_3	335–390
—— On the overconvergence of sequences of polynomials of best approximation	32_4	794–816
—— Note on the overconvergence of sequences of polynomials of best approximation	33_2	370–388
—— The existence of rational functions of best approximation	33_3	668–689
—— On interpolation and approximation by rational functions with preassigned poles	34_1	22– 74
—— Interpolation and functions analytic interior to the unit circle	34_3	523–556
WALSH, J. L., and RUSSELL, H. G. On the convergence and overconvergence of sequences of polynomials of best simultaneous approximation to several functions analytic in distinct regions	36_1	13– 28
WARD, L. E. Series associated with certain irregular third-order boundary value problems	32_3	544–557
—— A third-order irregular boundary value problem and the associated series	34_2	417–434
WARD, M. Postulates for the inverse operations in a group	32_3	520–526
—— The characteristic number of a sequence of integers satisfying a linear recursion relation	33_1	153–165
—— The distribution of residues in a sequence satisfying a linear recursion relation	33_1	166–190
—— Conditions for the solubility of the diophantine equation $x^2-My^2=-1$	33_3	712–718
—— The cancellation law in the theory of congruences to a double modulus	35_1	254–260
—— The arithmetical theory of linear recurring series	35_3	600–628
—— An enumerative problem in the arithmetic of linear recurring series	37_3	435–440
—— The diophantine equation $X^2-DY^2=Z^M$	38_3	447–457
WARSCHAWSKI, S. E. On the higher derivatives at the boundary in conformal mapping	38_2	310–340
WEATHERBURN, C. E. On curvilinear congruences	31_1	117–132
WEBBER, G. C. Waring's problem for cubic functions	36_3	493–510
WEGNER, K. W., and INGRAHAM, M. H. The equivalence of pairs of Hermitian matrices	38_1	145–162
WEISNER, L. Abstract theory of inversion of finite series	38_3	474–484
—— Some properties of prime-power groups	38_3	485–492
WEISS, M. J. The limit of transitivity of a substitution group	32_2	262–283
WHITEHEAD, J. H. C. On linear connections	33_1	191–209

	Vol.	Pages
Whitehead, J. H. C., and Lefschetz, S. On analytical complexes	35_2	510–517
Whitney, H. Non-separable and planar graphs	34_2	339–362
—— A characterization of the closed 2-cell	35_1	261–273
—— Analytic extensions of differentiable functions defined in closed sets	36_1	63– 89
—— Differentiable functions defined in closed sets. I	36_2	369–387
—— Differentiable functions defined in arbitrary subsets of Euclidean space	40_2	309–317
Whyburn, G. T. Cut points of connected sets and of continua	32_1	147–154
—— On the structure of connected and connected im kleinen point sets	32_4	926–943
—— Non-separated cuttings of connected point sets	33_2	444–454
Whyburn, W. M. Functional properties of the solutions of differential systems	32_3	502–508
Widder, D. V. On the expansion of analytic functions of the complex variable in generalized Taylor's series	31_1	43– 52
—— A generalization of Dirichlet's series and of Laplace's integrals by means of a Stieltjes integral	31_4	694–743
—— Necessary and sufficient conditions for the representation of a function as a Laplace integral	33_4	851–892
—— The inversion of the Laplace integral and the related moment problem	36_1	107–200
—— A classification of generating functions	39_2	244–298
Wiener, N., and Paley, R. E. A. C. Notes on the theory and application of Fourier transforms. I–II	35_2	348–355
—— Notes on the theory and application of Fourier transforms. III, IV, V, VI, VII	35_4	761–791
Wiener, N., and Young, R. C. The total variation of $g(x+h)-g(x)$	35_1	327–340
Wilder, R. L. Concerning zero-dimensional sets in euclidean space	31_2	345–359
—— A converse of the Jordan-Brouwer separation theorem in three dimensions	32_4	632–657
Wilson, W. A. Certain problems related to the cutting of a simply connected plane region by a continuum	31_3	552–562
—— On linear upper semi-continuous collections of bounded continua	31_4	829–836
—— On cyclic numbers of one-dimensional compact sets	34_2	263–273
Wintner, A., and Jessen, B. Distribution functions and the Riemann zeta function	38_1	48– 88

	Vol.	Pages
Wong, B. C. Certain contact properties of linear systems of hypersurfaces	37_2	207–215
Young, R. C., and Wiener, N. The total variation of $g(x+h)-g(x)$	35_1	327–340
Zippin, L. A study of continuous curves and their relation to the Janiszewski-Mullikin theorem	31_4	744–770
—— The Moore-Kline problem	34_3	705–721
—— Correction to a paper on the Moore-Kline problem	35_4	972
Zippin, L., and Montgomery, D. Periodic one-parameter groups in three-space	40_1	24– 36
Zygmund, A. On lacunary trigonometric series	34_3	435–446
—— Some points in the theory of trigonometric and power series	36_3	586–617
Errata, volume 26 (Wedderburn, J. H. M. Algebras which do not possess a finite basis, pp. 395–426)	32_4	944
Errata, volume 29 (Ritt, J. F. A factorization theory for functions $\sum_{i=1}^{n} a_i e^{\alpha_i x}$, pp. 584–596)	34_4	938